UNIVERSAL
LIBRARY

OU_160777

UNIVERSAL
LIBRARY

OUP—850—5-8-74—10,000.

OSMANIA UNIVERSITY LIBRARY

Call No. 330 B65P Accession No. G.8932

Author Blodgett, R.H.

Title Principles of Economics.

This book should be returned on or before the date last marked below.

PRINCIPLES of ECONOMICS

Revised Edition

RALPH H. BLODGETT, Ph.D.
Professor of Economics
University of Illinois

RINEHART & COMPANY, Inc.
Publishers *New York*

COPYRIGHT, 1941, 1946, BY RALPH H. BLODGETT
PRINTED IN THE UNITED STATES OF AMERICA
BY J. J. LITTLE AND IVES COMPANY, NEW YORK
ALL RIGHTS RESERVED

Preface

The present volume is such a complete revision of my *Principles of Economics,* published in 1941, that it might almost be described as a new book. It represents, as every textbook must, a compromise between two desirable but often conflicting objectives—simplicity and accuracy or completeness. I have not made simplicity an unconditional objective to be attained at all costs, but I have tried to present the discussion in as simple terms as possible within the limits set by the desire for accuracy or completeness. The result is a book which is probably not suitable for light, summer reading in a hammock, but it should be within the grasp of any average student who can achieve a moderate level of sustained effort.

I have tried to keep the present edition from having any aspects of a history of economic thought or of a survey of contemporary economic theory. Where conflicting definitions of terms, statements of principles, or theories exist, I have presented in general only the one which seems to me most suitable for beginning students. In presenting theory, moreover, I have endeavored in the present edition to state in each case the underlying assumptions under which the results predicted by theories or principles would be perfectly attained, and to indicate the extent to which these assumed conditions are likely to exist in our own economic system. As a result there are frequent cross references between our actual economic system and the ideal, theoretical model. Economic theories and principles are not presented as completely accurate descriptions of economic reality.

Many textbooks in economic principles confine themselves so rigidly to a description of the economic system of the United States that students are likely to get the impression that our capitalistic type is the only type and that our ways of doing things are the only ways. To ignore the fact that noncapitalistic systems are or have been in operation in various countries seems to me a lamentable oversight. While the present volume assumes that the primary interest of the American student is in his own economic system, nevertheless it gives him some idea of the institutions under which other systems operate and of the methods and practices employed in these systems. Again, in the theoretical sections, laws and principles are examined to see whether they have significance only in a capitalistic system or general validity for all types of economic systems. I hope that this approach will help to make the student's knowledge of economics both more sound and more useful than it would otherwise be.

Since this is a text on the *principles* of economics, it cannot undertake an exhaustive treatment of economic *problems*. Yet the dividing line between economic principles and problems is not sharp, and students who will take no other course in economics should not remain entirely unacquainted with leading economic problems of the day. Consequently the present edition contains chapters dealing with public finance and business fluctuations and employment, and sections devoted to monopoly, the influence of labor organizations and legislation on wages, obstacles to international trade, governmental regulation of banking, and other matters.

The order of topics and chapters in *Principles of Economics* is not entirely conventional, but I believe it is a logical and useful order. The first four chapters present introductory and background material. The next five chapters on production contain some important economic principles but are largely descriptive of our economic system, indicating the specialized, roundabout, and large-scale character of production, and the need for a tremendous volume of exchanges. A special effort in the present edition has been made to establish the connection between the principles of production and later theories of value and distribution. The significance of changing proportions of the productive agents, and of the principle of diminishing productivity, in connection with average, marginal, fixed and variable costs of production is shown in an introductory fashion in Chapter VI. The principle of diminishing productivity is stated and illustrated in such a way that it obviously becomes the basis of the theory of demand for any productive agent in the later theory of income distribution. The importance of the changing size of a firm in relation to cost is indicated in Chapter VII as an introduction to the later discussion of cost of production and value in the long run. The result will be, it is hoped, that the theories of value and distribution will not appear altogether new and strange when the student finally reaches them. In addition to abandoning the conventional discussion of the "point of most profitable use" in favor of a preliminary analysis of various types of cost in relation to price, I have brought together in a single principle of diminishing productivity the principles of diminishing returns and variable proportions, often stated separately. The analysis of size of firm in relation to efficiency and cost has been made to depend on the "indivisibility" of certain productive agents and services and the diseconomies of large-scale management. All discussion of exchange as such has been eliminated from the chapters on production.

The book then goes immediately into a discussion of the factors determining the rates at which products exchange. The individual topics in the theory of value are taken up in logical order in the present edition with no attempt to present all the easy topics first. After an introductory chapter dealing with the nature and significance of value and price and of various conditions of the market (competition, monopoly, duopoly, oligopoly, and

monopolistic competition), Chapter XI sets forth all matters pertaining to the theory of demand and the subject of marginal revenue. Chapter XII deals with all phases of cost of production and supply in various periods of time, and with the theory of opportunity cost, but rigidly excludes all matters of price determination and the interaction of demand and supply. Chapters XIII and XIV discuss the determination of price under competition in all periods of time, while Chapters XV and XVI do the same thing for the noncompetitive conditions of the market. In this new edition the chapters most thoroughly revised are those on value. Much of the value material is relatively new and some of it, though elementary, is not usually found in introductory texts.

Since distribution theory is essentially similar to value theory in that it studies the same pricing process from another point of view, the section on distribution follows that on value. This section includes an introductory chapter, which sets forth the general nature of the distribution analysis, the assumptions on which it is based, and its relation to earlier parts of the text; chapters on rent, interest, wages, and profits; and a final chapter on personal and noncapitalistic distribution.

The book then considers the mediums of exchange—money and credit. At the suggestion of some users of the first edition, this section has been cut from five chapters to three—one on money and two on credit and banking. These are followed by chapters on foreign exchange and international trade. The approach here is in terms of the newer total-cost analysis rather than in those of the traditional labor-cost analysis and in the present edition includes discussions of interregional trade, the tariff and other obstacles to trade, and the "Buy at Home" theory.

In view of the great and growing importance of governmental functions and expenditures, taxation, borrowing, and the public debt in our economy, Chapter XXVIII on public finance now seems appropriate even in a book devoted to the discussion of economic principles, and some discussion of fiscal policy seems desirable before passing on to an analysis of business fluctuations and unemployment. The traditional chapters on price levels and business cycles have been abandoned in favor of two chapters on economic instability which include, among others, sections on the validity of the concept of "business cycles," the underinvestment theory of depression, an analysis and criticism of the popular theory of economic maturity, the underinvestment theory and the great depression, the nature of full employment, and the need for and implications of a full-employment policy on the part of the federal government. The book concludes with a chapter on economic principles and problems in war and in peace.

Questions and problems are presented at the end of each chapter, and lists of selected outside readings accompany each chapter or related group of chapters.

There is no sterner discipline than the give and take of the classroom. The arrangement and content of the present volume have been developed to a great extent from almost twenty years of teaching experience, and also from the book's use in a preprinted form during the past two semesters in the large Economics 2 course at the University of Illinois. I am also obligated to many individuals: to all who read and criticized the manuscript and made suggestions for its improvement; to members of the Economics 2 staff at Illinois, who have made valuable criticisms and suggestions as a result of their using the book in its preprinted form; and to Professor Merlin H. Hunter, Chairman of the Economics Department University of Illinois, who has arranged my teaching schedules so that I would have an opportunity for investigation and writing, and has given me constant encouragement.

R. H. B.

Champaign, Illinois
May, 1946

Contents

I. THE NATURE OF ECONOMICS 3

Human Wants—The Means of Production—Economics as a Science

II. SOME BASIC ECONOMIC CONCEPTS 22

Goods, Economic Goods, and Wealth—Land, Capital, and Consumers' Goods—Income

III. ECONOMIC INSTITUTIONS 38

Institutions of Capitalism—Institutions of Socialism—Institutions of Communism—Institutions of Fascism

IV. THE PRICE SYSTEM 58

The Functions of the Price System under Capitalism—Obstacles to Equilibrium—Equilibrium and Optimum—The Price System in Noncapitalistic Economies

V. CHARACTERISTICS OF PRODUCTION 79

Production in General—Direct and Indirect Production—Specialization in Production

VI. PROPORTIONS OF THE PRODUCTIVE AGENTS 107

The Law of Diminishing Productivity—Diminishing Productivity and Costs of Production

VII. THE SIZE OF THE PRODUCTIVE UNIT 128

Size of the Productive Unit in Relation to Cost—Combinations of Productive Units—Monopolies

VIII. THE FORM OF THE PRODUCTIVE UNIT 151

The Single Proprietorship—The Partnership—The Corporation—The Cooperative Enterprise—The Governmental Enterprise

IX. RISK IN ECONOMIC ACTIVITY 173

The Nature of Risk—Insurance—Speculation—Hedging

X. INTRODUCTION TO PRICE DETERMINATION 194

The General Significance of Prices—Value and Price—Conditions of the Market

XI. THE NATURE OF DEMAND 211

Demand Defined and Illustrated—The Law of Demand—Demand and Market Demand—Elastic and Inelastic Demand—Demand for the Product of One Seller—Marginal Revenue

XII. SUPPLY AND COST OF PRODUCTION

Supply in the Short Run—Cost of Production and Supply in the Intermediate Period—Cost of Production and Supply in the Long Run

XIII. THE DETERMINATION OF PRICES UNDER COMPETITION

Price Determination in the Short Run—Price Determination in the Intermediate Period

XIV. THE DETERMINATION OF PRICES UNDER COMPETITION (*Continued*)

Price Determination in the Long Run—Interdependent Prices

XV. THE DETERMINATION OF PRICES UNDER NON-COMPETITIVE CONDITIONS

Short-Run Prices under Monopoly—Monopoly Price in the Intermediate Period—Monopoly Price in the Long Run—Class Price—Prices under Duopoly and Oligopoly

XVI. THE DETERMINATION OF PRICES UNDER NON-COMPETITIVE CONDITIONS (*Continued*)

Short-Run Prices under Monopolistic Competition—Intermediate-Period Prices under Monopolistic Competition—Long-Run Prices under Monopolistic Competition—The Validity of the Theory of Value in Noncapitalistic Systems

XVII. INTRODUCTION TO DISTRIBUTION

The Demand for the Agents of Production—The Supply of the Agents of Production—Assumptions of the Theory of Distribution

XVIII. RENT

Rent Determination for a Single Grade of Land—Rent Determination for Several Grades of Land—Other Considerations

XIX. INTERNATIONAL 359

The Nature of Interest—Interest Determination in the Long Run—Interest, Savings, and Investment in Short Periods of Time

XX. WAGES 376

Wage Determination for a Single Grade of Labor—The Validity of the Competitive Assumptions—Gross Wages and Net Wages—Differences in Wages between Groups of Workers

XXI. PROFITS 398

The Nature of Profits—The Causes of Profits under Competition—Profits under Monopoly and Monopolistic Competition

XXII. PERSONAL AND NONCAPITALISTIC DISTRIBUTION 412

The Personal or Family Distribution of Income—Noncapitalistic Distribution

XXIII. MONEY 433

The Nature and Functions of Money—Coinage and Currency—Paper Money—The Gold Standard—The Present Monetary System of the United States—Kinds of Money in the United States—Money and Economic Systems

XXIV. BANKING 457

The Nature of Credit—Investment Banking—Commercial Banking—The National Banking System—Defects of the National Banking System

XXV. BANKING (*Continued*) 482

The Structure of the Federal Reserve System—Elasticity of Currency and Credit—The Control of Commercial Credit—Safety for Depositors—Credit and Economic Systems

XXVI. FOREIGN EXCHANGE 508

The Uses of Foreign Exchange—Rates of Exchange between Countries on the Gold Standard—Rates of Exchange between Countries off the Gold Standard

XXVII. INTERNATIONAL TRADE 526

The Theory of Trade—The Balance of Trade—Obstacles to Trade—The International Trade of the United States

XXVIII. PUBLIC FINANCE 552
Public Expenditures—Taxation—Governmental Borrowing and the Public Debt

XXIX. ECONOMIC INSTABILITY 586
The Effects of Booms and Depressions—Business Cycles or No Business Cycles—Underinvestment and the Theory of Economic Maturity—Criticisms of the Theory of Economic Maturity

XXX. ECONOMIC INSTABILITY (*Continued*) 611
The Underinvestment Theory and the Great Depression—The Quest for Full Employment

XXXI. ECONOMICS IN WAR AND IN PEACE 635
Economic Principles in Wartime—Economic Problems in Wartime

INDEX 651

List of Figures

1. Curves of Average and Marginal Product 111
2. Average Cost and Average Variable Cost 120
3. Average and Marginal Costs 123
4. Average Cost Curves in Relation to Size of Firm 134
5. The Demand for Granulated Cane Sugar in the Market of X on a Given Day 213
6. Original, Increased, and Decreased Demand for Granulated Cane Sugar in the Market of X 214
7. Elastic and Inelastic Demand, and Demand Which Is Both Elastic and Inelastic 225
8. Demand for the Product of a Single Seller under Competitive Conditions 227
9. Demand and Marginal Revenue for a Monopolized Product 230
10. Original, Increased, and Decreased Supply of Granulated Cane Sugar in the Market of X 237
11. Average Variable, Average, and Marginal Costs for the Individual Firm in the Intermediate Period 241
12. The Supply Curve for a Whole Competitive Industry in the Intermediate Period 245
13. Average Cost Curves in Relation to Size of Firm, and Long-Run Average Cost Curve for the Individual Firm 247
14. Long-Run Average and Marginal Cost Curves for the Individual Firm 248
15. The Determination of the Price of Granulated Cane Sugar under Competitive Conditions in the Market of X on a Given Day 257
16. Effect of Increased and Decreased Demand on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions 260

17. Effect of Increased and Decreased Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions 261

18. Effect of a Simultaneous Increase in Demand and Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions 262

19. Effect of a Simultaneous Decrease in Demand and Increase in Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions 263

20. Short-Run Competitive Price Determination, under Conditions of Fixed Supply, in the Market of X on a Given Day 264

21. Price Determination for the Individual Firm in the Intermediate Period under Competition 267

22. Price Determination for a Whole Industry in the Intermediate Period under Competition 269

23. Long-Run Price Determination in a Competitive Industry under Conditions of Constant Cost 277

24. Long-Run Price Determination in a Competitive Industry under Conditions of Increasing Cost 280

25. Long-Run Price Determination in a Competitive Industry under Conditions of Decreasing Cost 282

26. Effect of Technological Changes on Long-Run Cost and Supply Curves under Competition 284

27. Price Determination in the Intermediate Period under Monopoly, Case I 297

28. Price Determination in the Intermediate Period under Monopoly, Case II 299

29. Price Determination in the Intermediate Period under Monopoly, Case III 300

30. Price Determination in the Long Run under Monopoly 303

31. Price Determination for the Individual Firm under Monopolistic Competition in the Intermediate Period 313

32. Rent Determination for a Single Grade of Land with Several Uses in the Long Run under Competitive Conditions 347

33. Productive Relationships among Four Grades of Cornland 351

34. Determination of the Net Interest Rate in the Long Run under Competitive Conditions 370

List of Tables

1. Results of Agricultural Production with Labor as the Variable Productive Agent 109
2. The Computation of Fixed and Variable Costs 118
3. The Computation of Average Cost 119
4. The Effect of Changed Prices of the Productive Agents on Average Cost 121
5. The Computation of Marginal Cost 122
6. Relative Importance of Various Branches of Economic Activity and the Percentage of Total Income Produced by Corporations in Each Branch, 1937 157
7. The Demand for Granulated Cane Sugar in the Market of X on a Given Day 212
8. Original, Increased, and Decreased Demand for Granulated Cane Sugar in the Market of X 214
9. Inelastic and Elastic Demand for Granulated Cane Sugar 223
10. Demand Which Is Elastic, Inelastic, and Unit Elastic for Granulated Cane Sugar 224
11. Demand and Marginal Revenue for a Monopolized Product 229
12. The Supply of Granulated Cane Sugar in the Market of X on a Given Day 236
13. Original, Increased, and Decreased Supply of Granulated Cane Sugar in the Market of X 237
14. Demand for and Supply of Granulated Cane Sugar in the Market of X on a Given Day 257
15. The Demand for Strawberries in the Market of X on a Given Day 263
16. Inelastic, Elastic, and Both Elastic and Inelastic Demands for Strawberries in the Market of X on a Given Day 294
17. A Summary of the Distribution of Income by Families in the United States in 1929 415

18. A Summary of the Distribution of Income by Families in the United States in 1935–1936 417

19. A Summary of the Personal Distribution of Income in the United States in 1944 417

20. Kinds and Quantities of Money in the United States, August 31, 1945 451

21. Securities Issued in the United States, 1927–1945 464

22. Balance of International Payments of the United States in 1940 542

23. International Trade of the United States, 1939–1945 548

24. Expenditures of the Federal Government, Fiscal Years 1941 and 1945 554

25. Sources of Federal Revenue, Fiscal Years 1941 and 1945 568

26. Expenditures, Revenues, Deficits, and Public Debt of the Federal Government, 1931–1945 580

27. National Income of the United States, 1929–1945 589

PRINCIPLES of ECONOMICS

I

The Nature of Economics

Economics is a very interesting subject for most students, for it deals with the business of living and making a living. All of us have certain amounts of money income, for example, and must make decisions as to how we shall use it. As a college student, you may have to decide on a given day whether to enjoy an elaborate dinner at a restaurant and spend the evening in your room, or get along with a less satisfactory meal and go to a movie. Should you use your last few dollars this week to pay your board bill, or take your girl friend to a big dance and hope to stall off your landlady or fraternity in the matter of your board? Should you secure your board and room rather cheaply at an ordinary boardinghouse, or will associations, friendships, and possibly pride of status justify the extra expense of belonging to a fraternity or sorority and living and eating at the "house"?

Should you buy all new textbooks and have less money left over to spend for other things, or get along with second-hand books and satisfy some of your other wants more completely? Should you buy a new winter overcoat this year, or make the old coat do another year and buy a new suit (or a couple of new dresses)? Should you work during the summer and save money which will enable you to live more comfortably during the rest of the college year, or should you go to college the year around on more meager resources in order to complete your education in a hurry? Should the co-ed invest in a new permanent wave or cover up her tresses in public with an attractive new hat? Problems of these kinds and many others have to be faced by most college students.

Later in life many other problems of income disposal may be encountered. Should a man try to acquire a desirable home of his own and be satisfied to drive an old-model car, or should he place his family and himself in cheaper, rented quarters and rejoice in a late-model automobile? Should the family have the latest type of radio and electric refrigerator even if all members must wear clothing that is a bit dowdy, or should more be invested in personal appearances even if radio and refrigerator have to be 1939 models? Should food consumption be lavish even if the members of

the family must relax after meals on old-style chairs and davenport, or would plainer table fare and more attractive furniture be more satisfying? Should the head of the family go to considerable expense to have doctors and dentists check up on himself and his family at regular intervals or should he spend his money on other things and hope for the best in matters of health? Should any part of the family's money income be saved and, if so, how much?

These and other economic choices with regard to the use of money income would not trouble an individual if his income were so large that he could readily satisfy his every desire, but most people have money incomes that are rather meager in comparison with the total prices of all the commodities and services which they would like to acquire and enjoy. The decisions which people make in disposing of their limited money incomes are very important both from the individual point of view and that of the operation of the economic system as a whole. That is, business enterprisers, in trying to decide on the commodities and services that are worth producing and the quantities in which they should be produced and offered for sale, are likely to be influenced strongly by the economic preferences and decisions of individual income-receivers or consumers.

In order that we may have money income to spend or save, many of us must work for wages or salaries. Even if an individual lacks specialized training which fits him for any particular occupation or profession, there may be quite a large range of jobs open to him and he must decide on the one which he will try to fill. He must decide to what extent a larger wage will compensate him for engaging in an occupation which is dangerous, unpleasant, or uncertain as to continuity of employment, and the extent to which the enjoyment which he receives from an occupation which is pleasant and attractive in itself will make up for relatively low money wages. He may have to decide whether it is worth while to move far away from home and live among strangers in order to secure a higher money wage; whether the higher wages obtainable in a certain occupation will compensate him for the increased efforts which will be required, in view of the increased taxes which will have to be paid out of the higher earnings; and whether he should seek an occupation in which it will be necessary for him to join a union or one in which he will deal as an individual with his employer.

If you are able to secure specialized training and education, should you prepare yourself to be a professional man, a salaried employee of a business or industrial firm, or a business enterpriser in your own right? How will you make a more specific choice in any one of these general fields? In attending college, how will you determine what part of your work should consist of courses aimed at general culture and development, and what part should be devoted to "bread-and-butter" courses which are supposed

to train you for making a living? How can you decide in advance what sort of occupation will be most likely to furnish you with a satisfactory income when you are finally ready to go to work?

When you come to earn an income and save part of it, should you invest your savings directly or put them into insurance or a savings account and let someone else worry about the task of final investment? If you will invest your own savings, will you put them into a business of your own or into an enterprise run by someone else? If you are going to purchase corporate securities, should you buy common stocks, preferred stocks, or bonds? How will you decide which industry should have the use of your savings and, more specifically, which firm in a particular industry? If you own land which you want to sell or rent, how will you decide who should be allowed to acquire it or have the use of it? These decisions with regard to labor, land, and savings are very important in the aggregate, for industries and enterprises cannot produce unless they have access to these so-called agents of production.

If you ever decide to operate an enterprise of your own, you will have many economic problems to solve. You must decide, on the basis of prospective prices of products and prospective costs of producing them, which commodity or service you should undertake to produce. You will need to determine how large your enterprise should be for greatest efficiency and net income; whether it should be organized as a single proprietorship, partnership, or corporation; how you will attempt to secure any necessary agents of production which you cannot furnish for yourself; and how you will attempt to deal with the various risks which attend the operation of any enterprise. You will have to decide on the proportions in which you will combine the various agents of production. For example, should you use a great deal of labor equipped with simple tools and devices, or less labor armed with complicated and expensive machinery and equipment? You will need to provide for the keeping of accurate records concerning your business, compute and pay taxes of various kinds to various governmental units, fill out numerous forms, reports, and questionnaires for the government, and be sure to comply with a large number of governmental regulations. These and many other problems make the life of a business enterpriser a complicated one.

Our discussion to date has indicated that each of us will be faced with numerous economic problems and decisions as life goes on. While economic problems are handled and economic decisions are made by millions of private individuals in a capitalistic system, the system as a whole operates in an orderly fashion. The desires of individuals for various commodities and services are expressed on the market by means of prices. Individuals decide to found and operate productive enterprises on the basis of the prices of commodities and services and the costs of producing them. Through these

decisions of enterprisers and their bidding in competition with each other, the limited available supplies of land, labor, and capital are distributed among industries and businesses, and the total quantities of capital and labor of specific types are increased or decreased. The end result is that the limited productive resources of society are used for the satisfaction of human wants. Our task in economics is to understand this process as a whole.

This task is not only interesting but important. From the study of economics you will come to have a better understanding of your own life and those of other people. You will know how the economic system as a whole operates. You will have a better chance to succeed in your own efforts to make a living and your attempts to use wisely the income which you acquire. And, finally, you will be a better citizen. From time to time you will have to vote for one candidate or another, or favor one public policy or another, and your study of economics will help you to make intelligent decisions. Should you vote for the candidate who favors relatively free trade between this country and other countries, or the one who espouses a high protective tariff and other obstacles to international trade? Should you vote for the candidate who favors the reduction of income taxes and an increased use of sales taxes and taxes on specific commodities, or one who would reverse these policies? Should oppressive private monopolies be disregarded, regulated by the government, or subjected to governmental ownership and operation? Should commercial bankers be left to their own devices in the matter of expanding and contracting credit, and in providing safety for their depositors, or should they be closely regulated by the government? Should full employment be made the primary objective of economic policy, and should the federal government pursue that objective regardless of consequences? To what extent is it the duty of government to provide for the economic security and welfare of the individual citizens? The principles of economics do not contain the answers to all such questions of policy, but they do afford a good basis from which to approach the problems.

HUMAN WANTS

On the basis of our introductory discussion we may define economics as *the science which deals with human wants and man's efforts to satisfy them through the use of scarce resources of production.* Human desires for commodities and services are clearly very great. Human wants have been expanding rapidly in the past and, as far as we know, are likely to continue to grow in the future. Some economists go so far as to say that human wants in the aggregate are unlimited. Others, more conservative, say that human wants, if not unlimited, are at least expanding without any known limit.

The Expansion of Human Wants. We may gain some idea of the extent to which human wants have been increasing in the past if, in imagi-

nation, we invade the modern, newly furnished home of a middle-class citizen and compare what we see there with the contents of the home of a person in a corresponding social position a century ago. Today we shall find that electric lights, and perhaps indirect lighting, have replaced tallow candles or kerosene lamps; a central heating plant burns coal, fuel oil, or gas, and supplies heat more efficiently than the old stoves and fireplaces; and overstuffed furniture is much more comfortable than the plain but durable furnishings of other days. We may find that the house is insulated to protect the inhabitants from the cold of winter and the heat of summer, or even that it is air-conditioned. We discover a telephone, a radio, a gas or electric range for cooking, an electric refrigerator, an electric washing machine, and numerous other electrical appliances, inner-spring mattresses, and a variety of beautiful and serviceable bathroom fixtures—all commodities which were not available to consumers a hundred years ago.

But our discoveries of modern commodities need not be limited to articles used in the home. We have paved, well-lighted streets and sidewalks; automobiles and airplanes for efficient, convenient transportation for business or pleasure; telegraph, cable, and radio service for communication; motion-picture theaters for entertainment; almost countless types of sporting goods for exercise and recreation; and expensive and elaborate equipment for the vicarious enjoyment of sports. It is also true that many of the types of machines and equipment in the factories and offices of today were not available many years ago, but our interest in these things is secondary at the present time because they contribute to the satisfaction of human wants indirectly rather than directly.

Many of the things which people consume today have partly or entirely replaced other commodities formerly consumed in large quantities. The radio has taken the place of the phonograph and the player piano in the home to a large extent; the development of the automobile caused a diminishing use of horses and carriages and, for a time, bicycles; and many men today wear soft collars already attached to their shirts instead of the stiff, detachable collars which once were popular. Although the problem of satisfying human wants is not what it would be if we had to produce the goods which people used to consume as well as those which are desired now, the desires of the average consumers for commodities and services today are considerably greater than they were some years ago and will undoubtedly continue to increase.

The Bases of Human Wants. For purposes of strictly economic analysis, our discussion of human wants might well stop at this point. In other words, in studying human behavior in so far as it has to do with the disposal of scarce means among alternative ends, we should consider the things which people want, or the ends which they desire to achieve, as given or ultimate data. Investigations of the reasons why people want the things they

do are usually undertaken in order to discover whether certain ends which people seek are "good," whether the realization of these ends would be conducive to "human welfare," or what methods of reform should be followed in order to "improve" the ends which people seek. These matters, as we shall see later, are no part of the business of the economist as a scientist. Accordingly, we shall leave the question of why people want various things to the biologist, the psychologist, the advertising expert, and the social reformer, for the most part. However, if we realize that the following discussion is in the nature of a digression and if we are careful to bear in mind that, for the purposes of economic analysis, the important thing is that people *have* certain wants and not *why* they have them, there can be little harm in examining a few suggestions as to the factors underlying human wants.

Human wants ordinarily have both a physical and a subjective side. On the physical side, they involve certain response mechanisms of nerves, muscles, and sense organs which stand ready to perform various functions when stimulated. On the subjective side, there is usually the consciousness that the use of the things wanted will afford pleasure or enable us to avoid the unpleasant. However, the subjective side of human wants is not always altogether trustworthy. We all know people who eat rather heartily but apologize for everything they eat, or fathers who announce that they attend the circus only for the purpose of affording enjoyment to their children. We all come into the world with very similar wants and we want originally comparatively few of the things which we later come to desire. A young baby, accustomed to a diet of milk, would readily starve to death amid the most luscious of beefsteaks. As time goes on, we are trained or conditioned to transfer our original wants to other things and we come eventually to have scores or hundreds of wants which were no part of our original equipment.

It follows that we must be cautious in classifying certain human wants as biological. To be sure, a certain amount of food is required to keep the individual alive and, in most parts of the world, certain minimum amounts of clothing and shelter are also necessary for survival. However, the desires of most people are not limited to the amounts and kinds of things which are necessary to survival. As a consumer of food the individual is concerned with satisfying his appetite rather than with remaining alive. Though adequate amounts of food could be found near at hand, the individual is likely to demand coffee from Brazil, tea from China, corned beef from Uruguay or the Argentine, cheese and wine from France or Italy, crab meat from the Far East, and a host of other things. Moreover, the consumers who live in northern climates are not content today, as their forefathers were years ago, to eat dried or preserved fruits and vegetables in winter months. They are likely to require fresh vegetables and fruits in

winter as well as in summer, even though these commodities must be brought in from warmer parts of the country or produced in local hothouses. In an up-to-date grocery store in the winter in some northern states, the consumer will find fresh carrots, green beans, peas, lettuce, spinach, broccoli, strawberries and other things which have been obtained from California, Texas, Louisiana, or some other southern state. Not only does he demand considerable variety in his diet, but the quantity of food which he consumes is seldom very accurately adjusted to his needs for existence or bodily efficiency. Further examples of the expansion of consumers' desires are scarcely necessary here for we would reach much the same conclusions if we examined the ordinary individual's wants in the matter of clothing and shelter.

We should not conclude, however, that individuals come in the end to have exactly similar though elaborate sets of wants for economic goods. Actually these wants differ considerably from person to person at any time, and from time to time for the same person. Among the factors which may be mentioned as conditioning the wants of the individual are his personal (hereditary) characteristics, his immediate physical environment, the income and social status of the family in which he is reared, his education, the occupation which he selects, the income which he is able to achieve, and many others.

On the other hand, there are likely to be at least broad similarities between the wants of different individuals. In this connection it is important to note that the individual consumes as a member of society, and not in isolation. An isolated individual's wants might well be limited to things which he deemed necessary for his physical existence and comfort, but the individual as a member of society is much influenced by his knowledge of what other individuals consume. Thus many people come to want certain economic goods when they see other people enjoying them and consider it a mark of social inferiority to walk while other people ride in automobiles, or to sit in silence at home while other people are listening to their radios. This tendency to match the consumption of other people, or to "keep up with the Joneses," is stimulated considerably by the tactics of advertisers and salesmen.

However, the individual is often not content merely to keep up with other people in the matter of consumption. Instead he may desire to outshine other people by consuming lavishly. Among the rich, this desire for ostentation may find expression through owning a residence as large as a medieval castle or by maintaining several residences for use at different times or seasons of the year, a fleet of expensive automobiles, a house full of servants, or a yacht which is an ocean liner in miniature. While this tendency to "conspicuous consumption," as it is sometimes called, is most obvious among the rich, a similar desire to excel other people in con-

sumption seems to exist in other classes of society. Even faculty bridge clubs have had to be broken up because of the desire of successive hostesses to outdo those who had gone before, so that evenings of entertainment which had started out with light refreshments, and trinkets for prizes, turned into lavish affairs with expensive prizes and elaborate refreshments.

The individual's consumption is also affected to a great extent by the fact that he produces on a specialized basis. If the individual tried to satisfy all his wants directly through his own efforts, his consumption would be confined to things for which he had already felt a desire and which he was able to produce. However, as we shall see in detail in Chapter V, the individual in our modern economic system specializes in a single line of production, or in some small part of a single line of production, receives a money income, and acquires through exchange the various commodities and services which he wishes to consume. In this way the range of commodities and services available for his consumption is greatly increased. Business enterprisers, whose main function is to provide people with the things they want, also have something to say about the things which people will want. For example, a business enterpriser frequently decides on the production of a new article, hoping that it will be profitable, and then tries through high-pressure advertising and salesmanship to convince consumers that they want his new product.

Finally, a part of the individual's consumption may depend on habit or custom. He becomes accustomed to the consumption of a certain type of good in satisfying a given want, and it is easier to keep on acquiring and using it than to make a change. Habitual consumption is not necessarily unintelligent consumption, for it rests on inferences made at some earlier time as to appropriate ways of getting what was wanted, and it may be a means of saving considerable time and effort providing the original inference was correct and the changing of the individual's situation has not made it obsolete. However, habits have a tendency to remain in force after the situation which gave rise to them has changed. To take an example from another field, mothers often caution their children to "look both ways" before crossing the street. The children may do so for a while, but they learn in time that, with traffic moving along the right side of the street, they need only to look to the left as they enter the street and later give a glance to the right as they get out into it. This habit may be satisfactory in the usual run of things, but many of our soldiers in World War II were hit by automobiles in Australia because they jumped out into the street after a glance to the left. The reason was, of course, that traffic in Australia moves along the left side of the street and, therefore, the first glance of anyone crossing the street should be to the right.

In summary, we see that human wants and human consumption are influenced by a large number of factors, but there is no doubt that human

wants as a whole are both great and growing. Again, though we have not discussed such an obvious matter, there is no doubt that individuals consider the satisfaction of some wants to be more important than that of others, and that individuals are able to arrange their wants in an order of preference.

THE MEANS OF PRODUCTION

The fact that human wants are large and growing would not result in economic activity—or the study of economics—if the means for the satisfaction of these wants were equally plentiful and growing at a similar rate. In the actual world in which we live, human wants are growing rapidly in the face of limited means for their satisfaction. In some parts of the world, it is true, human beings can exist largely on the bounty of nature with a minimum of effort on their part, but in most places the needs which can be satisfied without economic activity are limited to those for such things as air, sunshine, and, in some cases, water. However, the essential fact is not that finished goods, ready for human consumption, are limited in quantity. Though automobiles and suits of clothes do not grow on trees, the fundamental scarcity is that of the means of production, commonly called agents or factors of production, and not that of finished products.

The Scarcity of Land. Land is the agent of production which is most strictly limited in quantity in relation to the demand for it. To the economist, land means more than mere earth or soil, and may include rivers and streams, mineral resources, natural vegetation, and animal life. Although the amount of land under the control of one country may increase or decrease from time to time, it is clear that the land supply for the world as a whole is fixed in quantity for all practical purposes. Irrigation, drainage, or the discovery of better methods of production may make it profitable to bring into use certain soils or mineral deposits which were formerly not worth using, but this is not quite the same thing as increasing the quantity of purely natural wealth. Some land may at times be removed from the effective supply through erosion by wind or water, but all these changes are likely to be unimportant when related to the total quantity of land. It should also be noted that the total quantity of land is not of uniform quality for economic purposes. The land which is used in any particular branch of productive activity is likely to be of several grades or qualities, and the better grades are often especially limited in quantity.

The limited available quantity of land in general and of superior grades of land in particular would not be any great obstacle to the satisfaction of human wants if the other agents of production, labor and capital, were perfectly substitutable for land. In this happy situation, the shortage of land could be overcome by using small quantities of land in combination with very large quantities of labor and capital. In actual practice, however, labor

and capital are highly imperfect substitutes for land. This means, as we shall see in detail in Chapter VI, that attempts to increase production by applying larger and larger amounts of labor and capital to a fixed amount of land of given quality, under constant methods of production, are certain sooner or later to bring less and less satisfactory results in the form of progressively diminishing amounts of product per unit of labor and capital applied. And, of course, this tendency to diminishing returns on land of given quality cannot be offset by bringing poorer grades of land into use in the same branch of production. Thus the scarcity of land in general and of the better grades of land in particular, along with the fact that other agents of production are imperfect substitutes for land, constitutes the chief stumbling block in connection with man's efforts to satisfy his growing wants.

The Scarcity of Labor. Labor, the human element in production, is another agent of production which is relatively scarce. For some purposes labor is considered as including both ordinary labor and that of the business enterpriser; for other purposes the labor of the business enterpriser is considered separately. In the present connection the more inclusive meaning will be used. The total number of potential workers in existence is fixed at any given time, but the number of these workers who offer themselves for employment and the amount of work that they see fit to do are capable of considerable variation even in relatively limited periods of time. Thus the supply of labor (*human effort which is expended for the purpose of acquiring income*) should not be considered as fixed in short periods of time. Over long periods of time, the total number of workers, unlike the total quantity of land, may be increased or decreased significantly, though the forces which govern changes in population, number of workers, and labor supply are by no means completely economic in character. The total quantity of labor available at any time, like that of land, is not of uniform quality for economic purposes. It includes many grades, and the grades with the higher productiveness are especially limited in amount, whether because of an inherent scarcity of natural abilities or because the economic and social environment does not permit the natural abilities of some people to develop and show themselves.

While the population and resulting labor supply may be increased over a period of time, it is doubtful how much such increases contribute to the solution of the problem of satisfying human wants. It is true, as someone has observed, that each individual born into the world brings with him not only a mouth to feed but also two hands with which to feed it, but, unless a growing population shows increased ingenuity in devising more efficient methods and processes of production, the new arrivals in the population are likely to contribute more to the sum total of human wants than to the means for satisfying them. Since the individuals who are born into the

world perforce fail to bring any land with them, an increasing labor supply will be likely to assist in the solution of the fundamental problem of economic activity only if the talents of the larger population are able to increase the efficiency of productive methods and machines to the extent that the increasing relative insufficiency of land is more than offset.

The Scarcity of Capital. Capital consists of the nonhuman instruments of production which are not found in nature. Such things as machines and factories are produced by the application of human labor to land (assisted, once the process is begun, by capital already in existence). At any one time the total amount of capital in existence is fixed. Over a period of time, the total amount of capital may be increased considerably, but, since capital consists of produced goods, its production is subject to the same general limitations as the production of goods ready for final consumption.[1] Moreover, starting with fixed total amounts of the agents of production, we can produce more capital goods than formerly only by transferring agents from the production of consumers' goods to the production of capital goods. Thus if society as a whole is to acquire more capital goods than formerly, it must go without, for the time being, certain amounts of consumers' goods which it might otherwise have had.

Are the Agents of Production Really Scarce? Some people, especially in periods of business depression, are inclined to question the economist's decision that the agents of production are scarce relative to the sum total of human wants and in support of their opinion they point to the existence of large quantities of idle land, labor, and capital in our economic system. Under such conditions, they say, it is possible to increase the production of any given kinds of economic goods without taking agents away from other lines of production, and they conclude that the continued existence of large amounts of idle agents of production requires a fundamental alteration in our whole attitude toward economic matters. Now it cannot be denied that large quantities of the agents of production do stand idle in our economic system from time to time, and that this failure to use the available agents in the face of a tremendous volume of unsatisfied human wants seems regrettable to most people. Nevertheless, the basic conclusion of the economist with regard to the scarcity of the agents of production apparently remains sound, because, in so far as we can judge, our available supplies of the agents, even if they were all used full time, would fall far short of producing enough commodities and services to satisfy all human wants.

The fundamental scarcity of the agents of production is seen most clearly in wartime. During the years of World War II, for example, it became quite apparent that all of our agents of production, even if fully employed,

[1] It is possible, of course, that the total savings of individuals and firms in an economy at any given time may exceed the amounts which the same or other individuals and firms desire to invest currently in capital goods, but capital goods themselves are always scarce relative to the sum total of human wants to be satisfied by their use.

would not suffice to give us all of the commodities and services necessary for all-out prosecution of the war and at the same time maintain the desired volume of commodities and services for civilian consumption. The shortage of the productive agents was so apparent in wartime because both the war demands and the usual civilian demands for goods found effective expression in the market. But even in times of peace the total human desires for commodities and services, if they could find expression in the market, would make the total available quantities of productive agents appear just as inadequate. Agents of production may stand idle from time to time in our economic system, but this result should never be attributed to an insufficiency of total human desires or to a superabundance of the productive agents.

Alternative Uses of the Agents of Production. It is often said that mere scarcity of the productive agents in the face of great and growing human wants would not suffice to produce distinctly economic activity if each agent of production, or grade of an agent, could be used to produce only a specific commodity or service. Actually, however, each agent of production and grade of an agent can be used in the production of a considerable number of different commodities and services. In this situation we can satisfy any given want to almost any desired extent, if we are willing to do without various amounts of other things. We could probably produce enough automobiles, for example, so that each family could have five cars, but this result could be achieved only by transferring to automobile production considerable quantities of the agents of production which would ordinarily be used for other purposes and by reducing our consumption of various other economic goods. In all cases, the cost of obtaining a given quantity of one economic good is found in the alternatives forgone, or in the quantities of other economic goods which the same agents of production could have turned out. The alternative uses of the agents of production and the nature of the cost of obtaining increased quantities of certain economic goods stand out clearly in wartime. Thus, to take a single example, we found during World War II that we could get enough tanks, trucks, jeeps, airplanes, and other machines for war only by greatly reducing or eliminating the production of automobiles, electric refrigerators, washing machines, and other durable consumers' goods which require the same types of materials, labor, and other productive agents. The alternative uses of the agents of production introduce that element of choice which is necessary if human behavior in disposing of scarce means among given ends is to constitute genuinely economic activity.

ECONOMICS AS A SCIENCE

We have had a glimpse of the subject matter of economics but we also need in this introductory section of our work a brief discussion of the

methods which economics uses in dealing with this subject matter and of the limits of economics as a science. In the first place, what is a science? Some people define science as a method of studying phenomena, while others consider it to be the results obtained through the application of scientific method. Our preference is for the latter concept and we define science as *knowledge regarding any one department of mind or matter, co-ordinated, arranged, and systematized.* Through repeated experiment and observation, the scientist aims to discover laws or generalizations which govern the relationships and interaction of the phenomena in a certain field.

It appears, then, that economics as a science cannot stop with the mere description of economic activity. It must go beyond the mere facts concerning what is done and discover how and why. The description of economic activity might be interesting, but it would leave us unable to answer a host of questions. Why do some people work hard and receive only low wages while others apparently do less work but are much better paid or receive large incomes without working at all? Why are the prices of some goods low while those of others are high? Why is business depressed at some times and booming at others? Why is the general price level high or low? Why are the productive units large in some branches of production and small in others? In order to be able to answer questions such as these, economics must analyze and explain economic activity and not merely describe it.

Economics as a Positive Science. On the other hand, economics as a science must stop with the description, analysis, and explanation of economic activity. It is a positive science which studies what is or what tends to be, and not a normative science which prescribes what ought to be. Economics is the study of human behavior in the disposal of scarce means among alternative ends, and it is not concerned with discovering whether these ends are high or low, good or bad, worthy or unworthy. It cannot decide how our ends should be changed in order to maximize the general welfare. Its function is to investigate economic life and discover truths about it and not to hand down rules of life by which all good men should abide. It should be able to provide us with information concerning the probable consequences of given courses of action, but it cannot tell us whether the ends aimed at by these courses of action are worth attaining.

All this does not mean that the ends which men seek should never be evaluated or that means of increasing the social welfare should not be sought. In fact, we do not mean to bar the economist as an individual from making value-judgments or seeking the means of social improvement. All we are saying is that economics *as a science* cannot determine what welfare is or give a scientific answer to ethical questions. Along this line, Keynes has said: "At the same time, the greatest value is attached to the

practical applications of economic science; and it is agreed that the economist ought himself to turn his attention to them—not, however, in his character as a pure economist, but rather as a social philosopher, who, because he is an economist, is in possession of the necessary theoretical knowledge. It is held that if this distinction is drawn, the social and ethical aspects of practical problems—which may be of vital importance—are less likely to be overlooked or subordinated. Nevertheless, economics itself is a science, not an art or a department of ethical enquiry." [2]

The Inductive Method in Economics. We have seen that economics, as a science, must attempt to discover laws or generalizations which govern the relationships and interaction of the phenomena in the field of economic activity. But how are such laws or generalizations to be obtained? The obvious answer would seem to be that the economist should make a direct study of the facts concerning economic activity and from these facts derive his general laws or explanations. As a matter of fact, however, this method, which may be called inductive, has not been the source of most economic laws or theories, and we may well spend a short time in discovering why this is true.

For one thing, in dealing with economic phenomena, the economist can only observe; he cannot carry on experiments under controlled conditions. The student of chemistry can place a quantity of iron filings in a test tube, cover them with a certain amount of hydrochloric acid under controlled conditions, and be fairly confident of being able to observe and measure the results accurately. The economist cannot place the consumer in a test tube, pour a solution of lowered prices over him, and measure the results. Instead, he must rely on his observations of the consumer in the ordinary business of life where he will react to lowered prices in some way or other, amid a welter of other influences.

In the second place, in the absence of possibilities of experimentation, there are many phases of economic activity "in which direct generalization from complex economic facts, without a prior appeal to underlying principles, is generally speaking in a high degree untrustworthy, and in which, if we depend upon inductive reasoning alone, our conclusions are more likely to be erroneous than not." [3] Economic phenomena are extremely complex. In many cases essential facts are deeply hidden, while confusing or misleading items lie near the surface. Economic effects often flow from a plurality of causes rather than from a single cause, and economic forces often produce a strange intermixture of effects. Economic effects are often widely separated in time or space from their causes, and the more distant effects are sometimes more important than those which disclose themselves imme-

[2] J. N. Keynes, *Scope and Method of Political Economy*. London: Macmillan & Company, Ltd., 1891, p. 13.
[3] *Ibid.*, p. 197.

diately or near at hand. In some cases there is no genuine cause-and-effect relationship between economic factors. Instead, there are situations in which several economic factors are simultaneously determined. Finally, the phenomena studied in economics are in many cases subject to relatively great and rapid change. Before the effects of one set of causes can be completely worked out, other important sets of causes, beginning to work in different directions, interfere with the accuracy of the economist's observations. As one writer puts it: "One is reminded of a baby chasing a ball which a mischievous adult continually intercepts and throws in some new direction. The unfortunate infant may indeed set off on what, at the time, is an appropriate course; but that course always becomes irrelevant before the object in view has been achieved." [4]

These difficulties indicate the need for caution in using any methods in economics, but they render the method of pure induction particularly helpless and dangerous. As Lionel Robbins said, in referring to individuals who attempt to rely exclusively on inductive or quantitative methods in studying economic activity, "not one single 'law' deserving of the name, not one quantitative generalization of permanent validity has emerged from their efforts. A certain amount of interesting statistical material. Many useful monographs on particular historical situations. But of 'concrete laws,' substantial uniformities of 'economic behaviour', not one—all the really interesting applications of modern statistical technique to economic enquiry have been carried out . . . by men who have been themselves adept in the intricacies of the 'orthodox' theoretical analysis." [5]

The "Deductive Method" in Economics. What, then, is the method by which economic laws and principles have been obtained? It is often called (especially by critics) the deductive method, although it is really a kind of synthesis of induction and deduction. The economist is interested not in masses of detail but in the general aspects of economic phenomena and their typical relationships. In order to get at these things, he attempts to simplify the situations with which he must deal by making certain hypothetical assumptions concerning the nature of economic phenomena and the institutional framework in which they exist. When such assumptions have been made, economic phenomena are seen to have common aspects and relationships from which the economist can deduce general principles and laws. But economic theory would not be worth much if it stopped here. The next step is to study the extent to which the assumptions which have been made are likely to be realized in actual economic life, and the final step is to develop the modifications of economic theories, laws, and conclusions which will be necessary when the conditions which have been

[4] Barbara Wootton, *Lament for Economics.* New York: Rinehart & Company, Inc., 1938, p. 75.

[5] Lionel Robbins, *An Essay on the Nature and Significance of Economic Science.* New York: The Macmillan Company, 1937, pp. 114–115.

assumed are not present in the real world or are only partly realized in practice.

For example, if the economist wants to study the factors affecting the determination of wages, he does not go out into the labor markets of the country and consider all the thousands of actual wage rates which prevail and the varying situations in which they are determined. If he did, he might get a great deal of detailed data on wages but it is doubtful that he would ever arrive at any general principles or theory of wage determination. Instead, he decides to study one grade of labor at a time, under the institutions of capitalism, and under certain assumptions concerning labor markets. The owners of the agent of production in question and those who demand it in production are supposed to have a fairly wide knowledge of the market conditions which concern that agent. There must be so many persons demanding and supplying that productive agent that no one demander or supplier can affect its price by his actions. The employers of the labor must compete actively for the available quantity of it and the owners of the labor must compete actively for the available opportunities for employment. There must be no organized attempts on the part of workers or employers to force wages up or keep them down. Labor and capital are supposed to be mobile as between occupations and places. There must be substantial equality of bargaining power as between employers and workers. The government must not interfere with the process by means of which wages are determined but rather must leave wage rates entirely at the mercy of the forces of demand and supply in the market. Finally, the enterprisers who demand the labor must sell their own final products under competitive conditions.

Under these assumptions, the economist can see how demand and supply forces would determine wages and can state his general principles or theory of wage determination in terms of marginal productivity or some other factor. If the economist's deductions are accurate, his principles of wage determination are perfectly valid for any market situation in which the conditions which have been assumed are actually present. However, since these general principles of wage determination merely explain how wages would be determined under idealized conditions, they ordinarily furnish only a rough approximation to reality. They are of value not as a complete and final explanation of actual economic life but as a basis for study. With the general principles in mind, it is possible to see how actual economic conditions differ from those which have been assumed and to allow for these departures. That is, the economist must turn promptly to an examination of actual economic conditions in labor markets to see the extent to which his assumptions are realized or fail to be realized. Finally, he must explain how and to what extent his general principles and theories have

to be modified to cover actual market situations in which his assumptions fail to be realized perfectly.

It should be made clear that this general method of analysis is completely deductive only at one stage; that is, when the general principles and theories are being developed on the basis of assumed conditions. The economist's assumptions concerning institutions, market conditions, and economic forces at work are based on observation rather than deduction. In developing laws and principles for a capitalistic economy, the economist assumes that the institution of private property exists and not that most objects of wealth are publicly owned. He might assume that individuals in their economic activities are motivated by love of work and hatred of earnings, but he does not. On the basis of the actual functioning of people in our economy, he assumes that they are primarily influenced by economic motivation, or the desire for gain. Some of the basic assumptions of the economist, such as the one which holds that the various wants of the individual are of different importance to him and can therefore be arranged in a certain order of preference, are scarcely open to question. Other assumptions, such as those which refer to conditions of the market, may be imperfectly realized in actual economic life, but they are always realized to some extent and have some bearing on reality.

In similar fashion the economist relies on observation and not on deduction in re-examining his assumptions and in determining how nearly they approximate the facts of economic life. These observations enable him to decide to what extent his assumptions need to be modified, or the extent to which he should make allowance for the effects of various disturbing factors in cases in which it is not feasible to modify the original assumptions. Finally, the economist relies on observation in testing, confirming, and illustrating the economic laws and principles which he has developed by the process of deduction. The confirmation of economic principles, however, ordinarily means the discovery of cases in which the principles apparently afford a satisfactory explanation of actual phenomena, and not the discovery of phenomena which, on their own power, would have made it possible to develop the laws and principles which were actually obtained deductively. Needless to say, the results of various inductive studies may be of assistance to the economist in following his method of analysis. That is, these studies may help the economist in deciding what assumptions should be made originally, in determining the extent to which the assumed conditions approximate the facts of economic life, in testing, illustrating, and confirming his conclusions, and in discovering the fields with respect to which the conclusions need to be modified or in which additional analysis is needed.

There are several pitfalls which lie in the way of the economist in using the method of analysis which we have just described. His deductions

on the basis of assumed conditions may be inadequate or inaccurate. He may not state or even realize all the assumptions which he is making in developing general laws or theories. When these laws and theories have been developed, he may become enamored of them and forget that they were based on an extensive set of assumed conditions. Thus he may come to regard them as exact explanations of economic reality instead of rough approximations which are in need of constant revision in their applications to economic affairs. All these difficulties must be avoided if economic analysis is to be of the greatest possible value. In the end we may hope that, by using the method of analysis which we have described, we may come closer to our goal of understanding actual economic life than we could by studying it directly in all its complexity. And we must not forget that the aim and purpose of our study is the interpretation of economic reality.

QUESTIONS AND PROBLEMS

1. On the basis of our definition of economics, what are the fundamental conditions which must prevail before human activities can be characterized as economic? Explain.

2. How would you prove to a skeptic that human wants have been increasing over the last four or five decades?

3. Why are human wants likely to continue to increase in the future?

4. Write a brief essay on "The Bases of Human Wants."

5. "The nature of human wants is profoundly affected by the fact that individuals consume as members of society and not in isolation." Do you agree? Explain.

6. "The scarcity of land is the chief limiting factor in economic activity." Explain.

7. "While the total supply of labor is capable of increasing through time, its increase is likely to contribute very little to the satisfaction of human wants in general." Explain.

8. "The scarcity of capital is derived from the scarcity of the primary factors of production, land and labor." Explain.

9. Why does the economist consider land, labor, and capital to be scarce in spite of the fact that large quantities of these agents of production sometimes stand idle in the United States?

10. "The fundamental scarcity of the agents of production and the fact that these agents have alternative uses may be seen most clearly in wartime" Explain.

11. What is meant by saying that economics is a science?

12. What are the limits of economics as a positive science? Explain.

13. How is the economist handicapped in using the straight inductive method in attempting to discover economic laws and generalizations? Explain.

14. Outline the method which economists use in developing economic laws and generalizations.

15. "Economic laws cannot be worth much since they result from the exclusive use of the deductive method." Show whether you agree.

REFERENCES FOR FURTHER READING

Cassel, G., *Theory of Social Economy*. New York: Harcourt, Brace and Company, 1932, pp. 3–11.

Keynes, J. N., *The Scope and Method of Political Economy*. London: Macmillan & Company, Ltd., 1891.

Leven, M., and Others, *America's Capacity to Consume*. Washington, D. C.: The Brookings Institution, 1934, Chaps. 6–10.

Marshall, Alfred, *Principles of Economics*. London: Macmillan & Company, Ltd., 8th ed., 1920, Book I, Chaps. 1–4 and Book III, Chaps. 1–2.

Pigou, A. C., *Economics of Welfare*. London: Macmillan & Company, Ltd., 4th ed., 1932, Chap. 1.

Robbins, Lionel, *An Essay on the Nature and Significance of Economic Science*. New York: The Macmillan Company, 1937, Chaps. 1–6.

Stigler, G. J., *Theory of Competitive Price*. New York: The Macmillan Company, 1942, Chap. 1.

Wootton, Barbara, *Lament for Economics*. New York: Rinehart & Company, Inc., 1938, Chaps. 1–3.

Wyand, C. S., *The Economics of Consumption*. New York: The Macmillan Company, 1937, Chaps. 1–15.

II

Some Basic Economic Concepts

Before going any further in our study of economics, we shall find it desirable to define as precisely as possible some of the concepts which are to be used frequently in later chapters. However, since it will not be convenient to define all of the concepts of economics in the present chapter, the process of definition will be continued at later points in our discussion.

Problems of Definition in Economics. It is true, of course, that the mere definition of terms does not carry us far toward our goal of interpreting and understanding economic activity, and it is possible to waste much time in the process of developing definitions. However, many of the mighty economic controversies of the past have shown that much more time can be wasted when the participants in economic discussions use the same words with different meanings. Definitions of terms necessarily involve classification; that is, they indicate the items or phenomena which fall inside and outside particular classifications. While many definitions are somewhat arbitrary and it may not always make very much difference which definition of a particular term is used, we should try to bring together in each concept or classification those things which have the closest affinity for each other from the point of view of economics. Our aim is to make the concepts which are basic for the study of economics as clear and definite as possible and to point out those distinctions between phenomena which are of primary economic importance.

The task of defining terms is especially difficult in economics. Comparatively few people are amateur physicists, chemists, or biologists, and there is relatively little danger that students will come to the study of such sciences with popular meanings already attached to certain terms, such as molecule, electron, force, or mass, which will conflict with the meanings science has assigned to these terms. On the other hand, since many economic terms, such as wealth, income, land, and capital, are often used in ordinary conversation, many students approach the study of economics with popular meanings for these terms already fixed in their minds. Thus the question always arises as to the extent to which the meanings which are as-

signed to such terms in economics should correspond to the meanings which they have in popular usage.

Some writers are inclined to make consistency with popular usage the leading objective in defining terms in economics and hold that, if we assign unusual meanings to our terms, we may ourselves fall into inconsistency by occasionally and unconsciously using the same terms with their popular meanings. However, while we admit that some advantage is to be found in adhering to the popular meanings of terms, we also know that some of these popular meanings are extremely vague and that some terms have two or more different meanings in popular use. For this reason we conclude that the desire to give terms their popular meanings must be subordinated, where necessary, to the desire to set up definitions which are both clear and appropriate for use in economic analysis. Even though we define our terms very carefully, the complexity of economic phenomena will ensure that we shall sometimes have no easy task in deciding just which items fall within particular classifications.

GOODS, ECONOMIC GOODS, AND WEALTH

Goods. The things which people desire for the satisfaction of their wants are usually called goods. *A good is therefore anything which has utility; that is, anything to which the individual attributes a power to satisfy some want or other.* Utility is not a physical attribute of a good, such as its size or shape, for the same article (this textbook, for example) may give greatly varying amounts of satisfaction to different individuals at the same time or to the same individual at different times. Utility is a subjective and relative concept. Indeed, it is often said that utility is merely a relationship between individuals and things—a relationship which is expressed by saying that the individuals want the things (though we may reflect that things would not ordinarily be wanted if the individuals did not expect to be able to derive some use from them). Goods include both material objects, such as articles of food or clothing, and nonmaterial items, such as the services of actors or barbers.

As our discussion of the nature of economics in Chapter I suggested, goods and utility are terms which have no ethical or moral implications in economics. To call a thing a good does not imply that its use is conducive to individual or social welfare, and the decision that a thing has want-satisfying power does not imply that the want in question is one which should be satisfied. Certain goods are desired by people to satisfy wants which might much better be left unsatisfied from some point of view of social welfare, though individuals would disagree thoroughly as to just which things should be placed in this class. From the point of view of the economist, a Bible has utility and is a good, but the same things are true of a quart of whiskey or a roulette wheel. An article does not have to

possess want-satisfying power for all or most individuals in order to be a good. A burglar's jimmy or a wooden leg would be useful only to a limited number of persons, but they would be called goods nevertheless.

Economic Goods. While all goods are alike in possessing utility, they may be divided into two groups on the basis of scarcity and transferability. *Goods which are so plentiful that the individual may help himself to them without stint and without charge are known as free goods.* Air, sunshine, and, under some conditions, water are likely to be found in this group. Some free goods are physically capable of being transferred from person to person, but no free goods are really transferable in the economic sense. That is, they are so plentiful that no one takes the trouble of owning them or would be willing to pay anything to get them. Such things, of course, concern us very little in the study of economics, since it is unnecessary to economize in using free goods.

Things which are useful, scarce, and transferable are known as economic goods. Ordinarily the sign and token of an economic good is the necessity of paying a price, or of giving something in exchange, in order to get a certain good. Economic goods are scarce, then, in the sense that they exist in such limited quantities that the individual cannot help himself to them freely and without charge. They need not be rare in any absolute sense. Wheat in times of severe depression is scarce and qualifies as an economic good because it continues to command a price, although it may be available in such large quantities relative to the effective demand that its price is low and unsatisfactory to the growers. On the other hand, as someone has said, mosquitoes in the month of February are rare in northern climates but they are not economically scarce, for the number available, however small, is likely to be large relative to the number desired. Things must also be transferable in order to be economic goods. The fact that economic goods command prices indicates that they are exchanged between persons, and this means that the things themselves, or title to them, can be transferred from one owner to another.

Wealth. Some economic goods, in addition to possessing the qualities which have been mentioned, are also material, and to these things the economist gives the name "wealth." A thing is material if it is concrete or tangible in nature, as is a chair, a table, or a book. Thus, anything is wealth if it is useful, scarce, material, and transferable, and wealth as an aggregate may be defined as *a stock of useful, scarce, material, transferable things in existence at a given point in time.* The wealth of the United States, as of the beginning of 1946, included, according to this concept, all land (including natural resources) which was used in agriculture or other extractive industries, or as sites for buildings; houses, factories, stores, warehouses, theaters, hotels, office buildings, public buildings, and other edifices; tools and machinery in the factories and on the farms, and all equipment and

fixtures used in offices, stores and other establishments, and by members of the various professions; raw materials at the producers', in storage, or at the factories; the stocks in trade of wholesalers, jobbers, retailers, and others engaged in the marketing process; consumable goods in the hands of the final consumers, such as food, clothing, furniture, automobiles, books, and sporting equipment; and all other scarce useful things which were both material and transferable.

We should note, however, that our definition of wealth excludes the abilities and talents of the individual citizens of the country and the various valuable services which these individuals produce. Neither abilities nor services are material, and abilities, at least, are nontransferable. A person with ability of a certain kind may, of course, give lessons or training to other people, but his ability cannot be transferred directly to other people. If ability could be so transferred, the passing of an exceptionally able and talented individual would not involve so great a loss to the country as it now does. Services do pass from the persons who render them to the purchasers, and sometimes, as in the case of professional baseball or football players, the title to a person's services may be transferred from one owner to another for a consideration, but the services remain nonmaterial in character. Hence, if we wish to think of wealth in the aggregate as a *stock* of things in existence at a given point in time, wealth cannot be defined to include services, for an individual or a country does not have a stock of services on hand at a given time. In fact, both abilities and services have almost no significance as of a given moment in time. They exercise their influence *through time*. They are very important in connection with *income* (a concept which will be discussed later in this chapter), but we cannot include them in the category of wealth. It is also true that attempts to measure the wealth of an individual or a country would be somewhat handicapped if these various nonmaterial items were included in wealth. It is difficult enough to measure the value of our material possessions, but the problem of evaluating a host of nonmaterial things would be even more severe.

All this should not be understood as saying that material possessions are more valuable and important than nonmaterial things such as services and abilities. If we were willing to think of wealth as including all things which contribute to the "welfare" of the country and its citizens, abilities and services would have to be given a high rank, and even from the economic point of view it is true that a mere summary of material possessions does not give a highly accurate picture of the status of the country. Although the abilities of individuals, since they are not transferable, cannot be bought and sold, they are most important. If the people of a country were deprived of most of their material possessions, but were left with their abilities and skills intact, they could, it may be argued, bring them-

selves back to their former economic status after a time. However, if they retained their present objects of wealth, but lost their abilities and skills, the country could go only downhill from an economic point of view. Many services which people render are also highly valuable. For example, the baseballs, bats, gloves, and uniforms used by Bill Nicholson of the Chicago Cubs are worth only a relatively small number of dollars, while his services as a baseball player may readily command a price of twenty thousand dollars per year from his employers. We exclude abilities and services from the category of wealth not because they are not valuable and important but because their inclusion might easily lead to a great deal of confusion. Moreover, in omitting abilities and services, our concept of wealth rather well accords with popular usage of the term wealth.

Is Money Wealth? The question as to whether money should be counted as wealth is somewhat troublesome. Many people, accustomed as they are to thinking of wealth in terms of money, would not hesitate to answer this question in the affirmative, but economists are inclined to think that much of our money should be excluded from the category of wealth. Money is scarce in that it does not exist in such large quantities that one can help oneself freely to it without giving anything in exchange. It is material, and most of us find that it is all too transferable. Therefore, if money is to be excluded from wealth, it must be on the ground that it lacks utility.

At this point students often object strenuously. It is bad enough, they say, not to include money as wealth in the first place, but to throw it out on the ground that it is not useful is just too much. Surely money, as a medium of exchange, is useful in facilitating the exchange of economic goods. How many kinds of usefulness must a thing possess before it may be said to have utility? The answer is that objects of wealth in general have two kinds of usefulness. If a man owns a suit of clothes, he can wear it and thus satisfy one of his wants as a consumer or, if he does not want to wear it, he can ordinarily exchange it for something else. It is useful both in consumption and in exchange. Paper money, on the other hand, is useful in exchange but not in consumption. It has exchange value but not use value. If a law were passed forbidding further exchanges in our economic system, the man with the suit of clothes would find that this article still had usefulness in consumption, but a man who was holding a quantity of paper money would discover what the economist means by saying that money lacks utility in the ordinary sense of that term. He could, as students sometimes suggest, use pieces of paper money as bookmarks or as wallpaper, but, apart from such rather nonsensical purposes, he would find that his money was no longer useful once its exchange power was gone.

Metallic money constitutes an exception to the general rule that money should not be counted as wealth. Our various coins do contain certain mate-

rials which can be extracted and which have the power to satisfy certain wants in consumption, although the value of these materials as commodities may differ from their value in the form of money. To be sure, metallic money is not ordinarily useful in consumption so long as it remains in the money form. The gold, silver, or other metal must be extracted and changed in form before it can be used in consumption, but a similar change in form is necessary in the case of many things which we count as wealth, such as deposits of coal in the ground or standing timber in the forest, before they can be finally consumed. However, paper money, which makes up by far the greater part of the money in actual use in this country, consists of claims on wealth (and on other economic goods) rather than of actual objects of wealth. It is useful in obtaining, exchanging, and measuring wealth but it is not wealth itself. Clearly, then, any measure of the wealth of the country which included all of our scarce, useful, material, transferable possessions, and also all of the paper money with which these objects can be claimed, would be guilty to some extent of double counting.

Other Claims on Wealth and Income. There are several other things which are like paper money in these respects. That is, many people think of them as wealth but they are in reality claims on wealth and on other nonmaterial economic goods. Stocks, bonds, and mortgages are usually recognized by the economist as falling in this class. Some economists go further and include such items as patents, copyrights, and good will, although these claims seem somewhat more vague and intangible than stocks, bonds, and mortgages. Other things, such as theater and railroad tickets, are obviously claims on income alone, since they entitle one only to the use of wealth and the receipt of certain direct personal services. In any case, it is clear that all these things are valuable only as they enable their owners to obtain wealth and other economic goods, and not in and of themselves. To count these claims as wealth, in addition to the actual objects of wealth or income which they represent, is to engage in flagrant double counting.

For example, suppose that an individual owns practically all of the stock in a certain corporation. When the time comes to make assessments for the general property tax in the community the tax assessor looks over the land, buildings, machinery, and materials of the corporation and sets a valuation on them which requires the payment of a certain tax. Then he wanders around to our individual's home, discovers his stock certificates representing the ownership of this corporation, and sets a valuation on this stock which necessitates the payment of another tax comparable in size to the first one. Our hypothetical corporation owner would probably decide at once that his wealth had been counted and taxed twice and that in the corporation and its stock he really owned a single set of objects of wealth

and not two sets. These conclusions, of course, are those of the economist also.

Some economists distinguish between the individual and the social points of view by saying that such things as stocks, bonds, and mortgages may be considered as wealth for the individual owner but not for society as a whole. If I own a mortgage on a man's farm and something happens which renders invalid my claim on his wealth, what has happened? It might be said that I have suffered a loss of wealth and that the man who gave me the mortgage on his farm has gained, but society as a whole has neither gained nor lost. The country has just as many farms as before, equipped with the same buildings and machinery, and with soil of the same fertility and productivity. Such arguments serve to emphasize the economist's point that, while such items as paper money, stocks, bonds, and mortgages are of considerable importance to various individuals, they should not be counted as part of the national wealth.

Measuring Wealth in Terms of Money. The wealth of a country is composed of so many different kinds of articles that its measurement in physical terms is very difficult. Since it is hardly feasible to add together directly such diverse items as bushels of wheat, automobile factories, golf courses, locomotives, and perambulators, we ordinarily fall back on money as a means of expressing the total wealth of the country. The measurement of the national wealth in terms of money is satisfactory enough at any given time, but, unless the period in question is one in which the general price level and value of money remain approximately constant, successive measurements over a period of years are likely to give a somewhat inaccurate picture of what is happening to the national wealth. When the general price level moves up and down, the value or purchasing power of money decreases and increases, and such changes in the value of money are reflected in estimates of the national wealth. Thus when we are told that the national wealth of the United States amounted to about 354 billions of dollars in 1929 but only to 287 billions of dollars in 1934, we should not assume that this country had at the latter time only about four fifths as many scarce, useful, material, transferable things as in 1929. In similar fashion, an increase in the estimate of national wealth from 287 billions of dollars in 1934 to 322 billions in 1937 did not necessarily reflect a comparable increase in our holdings of the physical objects of wealth.[1]

LAND, CAPITAL, AND CONSUMERS' GOODS

Land. For purposes of analysis, it is desirable to divide the wide range of articles called wealth into three parts—land, capital, and consumers' goods. By land we shall mean *all natural wealth; that is, wealth which*

[1] National Industrial Conference Board, *Enterprise and Social Progress.* New York: 1939, p. 58.

exists without human labor having been applied to produce it. It includes a large part of the earth's soil, with its qualities of fertility, topography, and climatic conditions as they exist in nature; natural vegetation such as trees, plants, and fruits; wild animals and fish; mineral deposits, such as petroleum, iron ore, and coal; and streams and bodies of water, in so far as all these things satisfy the requirements of wealth of being scarce, useful, material, and transferable.

In deciding what should or should not be included in the category of land, it should be remembered that land must, above all, be wealth. It must also be natural, but natural is only the adjective while wealth is the noun. The student should avoid classifying anything as land merely because it exists in nature and without regard to whether it is scarce, useful, material, and transferable. Thus any part of the earth's soil which is so poor in quality that it is of no use to anyone, or which exists in such large quantities that it does not command a price, should be excluded from the class of things called land. The same considerations, of course, apply to natural vegetation and animal life, mineral deposits, and streams or bodies of water. We should also avoid the mistake of classifying such things as fertility, sunshine, and temperature as land. These things may be ever so natural, but they are merely characteristics which go along with the land and are not land itself.

It should be clear from the definition of land that very little land exists in the pure state in any well-developed economic system. That is, most of what was originally land will have had human labor and capital applied to it in such a way that it is no longer possible to separate out the natural elements. Our definition of land is also somewhat at variance with common or business terminology. Businessmen quite commonly consider their holdings of land as part of the capital with which they do business, and they do not attempt to distinguish between the natural and the man-made instruments of production. However, it cannot be denied that land, in the sense of natural wealth, is the foundation upon which the entire superstructure of economic activity is erected, and it will often be desirable in our later work to have a concept of land which is distinct from that of capital.

Capital and Consumers' Goods. Capital is frequently defined as produced goods which assist in further production, while consumers' goods are said to be finished products in the hands of the final consumer. These definitions may seem satisfactory at first glance, but a close examination will reveal that they do not distinguish at all sharply and accurately between these two classes of wealth. For example, if we are asked to classify coal which is burned in the factory as either capital or consumers' goods, our first inclination is to say that it is capital, since it is a produced good which is used in further production. However, it could also be classed as a consumers' good without straining the definition very much, for it is a fin-

ished product in the sense that it will never be subjected to any further processing, and it is in the hands of the final consumer, since no one will use the coal as such after it has been consumed in the factory. In similar fashion, we should classify food on the workingman's dinner table as a consumers' good, since it is a finished product in the hands of the final consumer. However, this food is a produced good and, if fit for human consumption, it should have the effect of assisting, indirectly at least, in further production.

In view of these difficulties, we shall define capital for our purposes as *produced wealth which assists directly in further production and which furnishes satisfaction or gratification to the user only indirectly or incidentally, if at all.* According to this definition, coal burned in the factory would undoubtedly be classified as capital. The burning of it may furnish heat which is enjoyed by the workers, but any pleasure which they derive is purely incidental to the main purpose in using the coal, which is to facilitate the processes of production. If the employees could work more efficiently in an unheated factory, it would doubtless remain unheated. Again, while an employee may derive some personal satisfaction from the operation of a locomotive, a delivery truck, or some other complicated piece of industrial or business equipment, we should have no difficulty in classifying these things as capital.

Consumers' goods, on the other hand, may be defined as *produced wealth which directly gratifies human wants in consumption and which assists in further production only indirectly or incidentally, if at all.* Thus food on the workingman's table should be classed as a consumers' good since the worker uses it to satisfy directly one of his wants as a consumer and since the fact that this food may assist him in further production the next day is purely incidental to this main purpose. People usually eat for enjoyment and not to make themselves efficient workers. In similar fashion, coal burned in the furnace at home would be classed as a consumers' good, unless the homeowner or renter takes roomers, because it operates to satisfy a basic want in consumption, and the fact that a worker with a well-heated home may be more efficient in production than one from an unheated home is purely incidental to this main purpose. In classifying things as consumers' goods, however, we should see to it that they are first of all wealth. Students sometimes think of motion-picture performances or doctors' services as consumers' goods, but they are not, although they are enjoyed by individuals as consumers. Since they are nonmaterial economic goods, or services, they are included in income but not in wealth or consumers' goods

According to these definitions, we see that the capital of a country includes its factory buildings, warehouses, stores, and all other buildings which are used for productive purposes; machinery, tools, and other equipment of factories, farms, stores, offices, and all other productive establishments;

raw materials and supplies in the hands of producers or of industrial users; goods in process and finished products at the producers'; the stocks in trade of retail stores and other merchandising establishments; and any equipment used by persons who render direct personal services. The student should not hesitate to classify as capital finished goods in the hands of wholesalers, jobbers, retailers, or other persons engaged in marketing, nor should it trouble him to count as capital the equipment and tools used by producers of services. As we shall see later on, production is defined broadly in economics to include the activities of people who transport, store, or market objects of wealth or who render direct personal services, as well as those of persons who "make things." Thus all persons who create utilities are producers, and all produced objects of wealth which assist them directly in performing their functions should be considered capital. The nation's consumers' goods, on the other hand, include all produced objects of wealth which are actually in use (and not merely on their way toward being used later on) for the direct satisfaction of human wants in consumption.[2]

Classifications of Capital. Those objects of wealth that are called capital may be subdivided in several ways. One such classification divides capital into *fixed* and *circulating* capital on the basis of durability. Fixed capital is that which is durable and which may be used repeatedly over a long period of time, while circulating capital is nondurable and is used up in a single process or a small number of processes. The furnaces of a factory, its machinery, and the typewriters in its office are fixed capital. The coal burned in the furnaces, the raw materials which pass through the machines, and the stationery used in the typewriters are examples of circulating capital. It is also possible to classify capital on the basis of the number of different purposes for which it may be used, as well as according to the number of times which it may be used for a given purpose. If a capital good can be used in many different kinds of production, it is called *free capital,* while if it can be used only in one line or in a very few lines of production, it is called *specialized capital.* Thus, coal, lubricating oil, and many kinds of raw materials, such as cotton or rubber, would be considered free capital because it is possible to use them in numerous industries. On the other

[2] Under any definitions, some difficulties remain to trouble us in distinguishing between capital and consumers' goods. Most people would classify electric refrigerators and washing machines in the home as consumers' goods. However, consumer demands for ice and laundry service not so many years ago would have been satisfied in large part by ice manufacturing plants and public laundries, and the plant and equipment of these productive establishments would then have been classified as capital. Some changes also occur in the other direction. The pans, bowls, cookstoves, and other household goods which were once used to provide the family with bread would have been classified as consumers' goods, but the plant and equipment of modern bakeries which now supply us with the same product are considered to be capital. As a result of such considerations, some economists propose that we should consider all durable goods as capital, but serious difficulties are also encountered when this approach is used.

hand, if a manufacturer has machinery for making automobiles, he cannot use it to turn out tennis balls or to bottle soft drinks; such capital would be called specialized.

Converting Land into Consumers' Goods. The definitions which have been given for land, capital, and consumers' goods suggest that it is frequently impossible to classify an article accurately unless we know how it is being used. A piano in one's home, for example, would usually be considered a consumers' good, but it may easily function as a piece of capital if its owner uses it to give music lessons. In similar fashion, even a rich man's yacht may serve as a capital good if, in a time of national emergency, he turns it over to the government for the transportation of freight. We also learn from a consideration of our definitions that it is quite common for a thing to begin its economic career as land, to pass through an intermediate stage when it must be considered capital, and to finish as a consumers' good. Thus a fish in the stream (unless planted there by human beings) would be called land. In the fish market, as part of the enterpriser's stock in trade, it is capital, and on the dining table at home it has become a consumers' good.

Iron ore in the ground is land, but iron ore, pig iron, or a steel billet at the steel manufacturer's is capital. The sheet steel in finished automobiles is still capital as long as the cars are in the hands of the manufacturer or dealer, but when the cars pass into the hands of individuals who drive them for pleasure, the steel and other materials in the cars are consumers' goods. However, the steel billets might just as well have been converted into textile machinery or equipment for dentists or barbers. In such cases, the steel would never reach the final stage of being consumers' goods, but would spend its useful life as capital, assisting, when combined with human labor, in the production of further articles of wealth or personal services.

INCOME

Wealth and Income. Income is a term whose importance in economic discussions corresponds closely to that of wealth. There are many definitions of income, but we prefer to define it as *a flow of economic goods over a given period of time.* Thus income and wealth are related concepts, but there are significant differences between them. Wealth is thought of as a stock or aggregation of things as of a given point in time, while income (also called *real income*) is considered as a flow of things over a period of time. The income of an individual during a month or a year is a meaningful concept, but it is impossible to think of the income of an individual as of a given point in time. In the second place, wealth consists only of those economic goods which are material, while income includes all new objects of wealth received in a given period and in addition any non-

material economic goods, such as services, which are received in the period.

Let us further clarify these concepts by considering the relationships which exist between the wealth and income of the nation in a given year, say 1946. At the beginning of the year the wealth of the nation consists of all the scarce, useful, material, and transferable things (divided into land, capital, and consumers' good) which it possesses. During the year labor is applied to the land and capital, and this process results in part in the production of new consumers' goods and new capital goods. Another part of the combination of human labor with land and capital results in the production of personal services, such as those of actors, doctors, lawyers, and barbers, rather than new objects of wealth, and these personal services, as well as the new capital and consumers' goods, are included in the income of the nation for 1946. But this is not the end of the story. The consumers' goods in the possession of the people at the beginning of the year continue to render valuable services to their owners during the period, and these present services of consumers' goods acquired in the past (that is, in preceding income periods) are a part of the national income in 1946. Objection is sometimes raised to the inclusion of the present services of previously acquired wealth in current income, but it appears logical to include these items. Quite apart from the ownership of these objects of wealth, the present services of, say, an automobile or a dress suit are useful, and they are scarce in that a price must be paid by a nonowner in order to obtain them. Since these services qualify as economic goods, and are commonly bought and sold, they should be included as income.

The present services of previously acquired wealth are, of course, used up during the income period, although the objects of wealth themselves may continue to yield services in future periods. Similarly, personal services are consumed completely within the income period. Of the new consumers' goods received as income by our people in 1946, part are nondurable and may be used up during the period, as in the case of food, gasoline, and some articles of clothing. Even such things, however, if they are not completely perishable, may be held over for later consumption. Of the more durable new consumers' goods received as income, some items may replace other consumers' goods which have become worn out during the period, while other items enter into and become a more or less permanent addition to the wealth of the nation. Much the same thing is true of the new capital goods received as income in 1946. Some of them, being nondurable in character, are likely to be used up during the period. Some of the new durable capital goods go to replace old capital goods which have become worn out during the period, and others constitute a net addition to the wealth of the nation as of the end of the year. Thus in the ordinary year, in spite of the need for replacements, the wealth of the nation

in terms of physical goods may be expected to be greater at the end of the year than at the beginning.

Real, Psychic, and Money Income. The income which we have been discussing, composed of new objects of wealth, personal services, and the present services of old wealth, received during the income period, is known as real income. In addition there are two other concepts of income which should be mentioned. One of these concepts, psychic income, refers to the enjoyments or satisfactions received by the individual over a given period of time. We refer to psychic income here only to call attention to the fact that such satisfactions or benefits constitute the final objective in the receipt of income. Otherwise the concept of psychic income is of comparatively little use. It certainly does not correspond closely with what is meant by income in everyday or business discussions. The national income which is distributed annually among the owners of the agents of production does not consist of psychic benefits, and, since the same good may furnish widely different amounts of satisfaction to different consumers at a given time or to the same consumer at different times, this concept of income would be impractical for purposes of measuring the national income and for income taxation. Finally, an individual's money income and real income could vary widely from his income in terms of benefits or satisfactions.

Money income, or the amount of money received over a given period of time, is a concept of income with which everyone is familiar. Of course money income is not really important in and of itself, because what people really want in seeking income is not money but the things money will buy. Money income would not be satisfactory as our only concept of income, because money income does not furnish a very reliable index to the amount of real income which is received. For one thing, some real income is received by individuals which does not register in anyone's money income, as in the case of food raised by the farmer and consumed on the farm, or personal services which one receives in the home from other members of the family. More important is the fact that one's money income may change considerably without any comparable change taking place in one's real income, because the value of money itself changes with changes in the general price level. Thus a man with a money income of three thousand dollars in a given year might enjoy a real income as great as, or greater than, that which he received in another year when his money income was four or five thousand dollars.

Since the real income of an individual or a nation is composed of such a great variety of items, it is often convenient to express income in terms of its money value, but we should remember that such money measures do not always give a completely accurate picture of what happens to real income in successive periods of time. Thus, when we read that the national income of the United States amounted to 83 billions of dollars

in 1929 but only to 40 billions of dollars in 1932,[3] we should not suppose that the flow of actual economic goods was only half as great in the latter year as in the former, for much of the apparent decline in income was due to a fall in the prices of the items included in real income rather than to a shrinkage in the actual quantities of these items. Similar considerations should be borne in mind in comparing an estimated national income of 71 billions of dollars for 1939 with one of 149.4 billions for 1943 or 160.7 billions for 1944.[4] Moreover, general estimates of the national income are unreliable for our purposes for another reason; that is, they include only the earnings of individuals in producing new wealth and personal services during the year and do not make allowances for the services of old wealth currently used in direct consumption. With such an allowance made, the national income for 1929 was about 92 billions of dollars,[5] and estimates for later years would also have to be increased by several billions of dollars in order to measure real income completely. However, in spite of the unreliability of money income as a measure of real income in individual years and from one year to another, the concept of money income is useful to some extent in economic analysis, and the taxation of incomes in the United States is based upon receipts of money or the equivalent thereof in terms of commodities and services.[6]

Scarcity Definitions. Before concluding this chapter we should note that most of the definitions which have been given (such as those of economic goods, wealth, land, capital, consumers' goods, and real income) are scarcity definitions; that is, they deal with things that exist in quantities which are quite limited relative to the desires of people for them. Such definitions are necessary and desirable, as things are, but we should not allow them to lead us to unsound conclusions.

Quantities of wealth and income are commonly expressed in terms of money, and a small quantity of wealth or income may have a total money value which is as great as, or greater than, that of a larger amount of wealth or income under other circumstances. Assuming no international trade, a country with a national wealth of 100 billions of dollars and a national income of 25 billions might have a higher average standard of living and possibly a higher level of economic welfare than that enjoyed in another country with a population of the same size and a national wealth and income twice as great in money value as those of the first country. This could happen if the small national wealth and income of the first country were due not to scarcity but to abundance, so that many things

[3] *Survey of Current Business.* Washington, D.C.: Department of Commerce, July, 1936, p. 2.

[4] *Survey of Current Business,* February, 1945, pp. 3–5.

[5] M. Leven, H. G. Moulton, and C. Warburton, *America's Capacity to Consume.* Washington, D.C.: The Brookings Institution, 1934, p. 148.

[6] Fluctuations in national income are of extreme importance to the individuals of the economy, and will be dealt with in some detail in Chapters XXIX and XXX.

which registered in the national wealth and income of the second country, as measured in money, were so plentiful in the first country that they had become free goods and thus had failed to be counted in wealth and income as defined in terms of scarcity and as measured in terms of money. By the same token, a country would reach the highest standard of living and possibly the highest level of economic welfare if all of the things which its people desired had become so plentiful that they were free goods and nothing remained which could be counted as wealth or income, as we have defined these concepts in terms of scarcity.

QUESTIONS AND PROBLEMS

1. "Consistency with popular usage should be the leading objective of the economist in defining terms in economics." Discuss.

2. Distinguish between free goods, economic goods, and wealth.

3. "Wealth is a stock of useful, scarce economic goods in existence at a given time." Do you agree? Explain.

4. "Utility is an essential characteristic of wealth and all tangible things that have utility are wealth." Comment.

5. "Since a worker's skill is just as important in his acquisition of income as are the capital instruments and land with which he works, there is no logical reason for saying that skill is not wealth while capital and land are wealth." Show whether you agree.

6. "Since it is desirable to have our economic concepts in agreement with everyday usage, all money should be included in the category of wealth." Criticize.

7. In deciding whether money should be counted as wealth, does it make any difference whether one is considering paper money or metallic money? Why?

8. What is the difference between wealth and claims on wealth? between individual wealth and social wealth?

9. "The tremendous increase in the bonds issued by the federal government in recent years has greatly increased the wealth of the United States." Do you agree? Explain.

10. "In order to be wealth, goods must be scarce; in order to be wealthy, therefore, a society should endeavor to make things scarce." Show whether you agree.

11. Explain the difficulties which are encountered in measuring the national wealth and income in terms of money.

12. "The concept of land in economics includes some things which are not ordinarily thought of as land and excludes other things which are commonly considered a part of land." Explain.

13. "The definition of capital as 'produced goods which assist in further production' and that of consumers' goods as 'finished products in the hands of the final consumers' furnish a satisfactory basis for distinguishing between capital and consumers' goods." Show whether you agree.

14. "Durable goods are capital but nondurable goods are consumers' goods." Do you agree? Why?

15. "Circulating capital differs from fixed capital in that it passes freely from hand to hand and changes possession frequently, while fixed capital is not so readily moved and changes possession less frequently." Show whether you agree.

16. "Under any definitions, some difficulties remain to trouble us in distinguishing between capital and consumers' goods." Explain.

17. Show how a given economic good may qualify at different times as land, capital, and consumers' goods.

18. Explain carefully how the concept of income differs from that of wealth.

19. "Since money income is useful only because it can be exchanged for real income, it follows that a person's money income and real income are equal and there is no need to distinguish between them." Do you agree? Explain.

20. Classify the following items as goods, economic goods, wealth, land, capital, consumers' goods, and income. Place each item in as many of these categories as possible, according to the definitions in the text, but do not classify items which do not belong in any of the categories.

a. Stock in the Ford Motor Company
b. A mug of beer
c. A U. S. Government bond
d. Furniture at the dealer's
e. A motion-picture performance
f. A ticket for a football game
g. A dollar bill
h. Your family automobile
i. The present services of your family automobile
j. Burglar's tools
k. A pair of beautiful blue eyes
l. Timber standing in the forest

REFERENCES FOR FURTHER READING

Cannan, Edwin, *Wealth*. London: P. S. King, 2nd ed., 1928, Chap. 28.

Carver, T. N., *The Distribution of Wealth*. New York: The Macmillan Company, 1913, Chap. 3.

Hobson, J. A., *Work and Wealth*. New York: The Macmillan Company, 1922, Chaps. 3–9.

Marshall, Alfred, *Principles of Economics*. London: Macmillan & Company, Ltd., 8th ed., 1920, Book II, Chaps. 1, 2, 4.

Pigou, A. C., *Economics of Welfare*. London: Macmillan & Company, Ltd., 4th ed., 1932, Part I, Chaps. 1–3.

Taussig, F. W., *Principles of Economics*. New York: The Macmillan Company, 4th ed., 1939, Chap. 1.

III

Economic Institutions

The general nature of the economic problem—that of satisfying the large and growing wants of people through the use of scarce means of production which have alternative uses—remains always the same, but the economic activity which results may be carried on under a number of different kinds of economic systems, such as capitalism, socialism, communism, or fascism. The forms which economic activity takes, and the methods and policies which are employed in carrying on this activity, may differ considerably from one type of economic system to another. They may be one thing in an economic system in which the activities of government are held to a minimum, and in which the institutions of freedom of enterprise, private property, and competition prevail, and quite another thing in a system in which the government owns and operates all industries and controls all economic activity, so that the institutions named above are completely absent. Similarly some of the economic laws or generalizations advanced as being descriptive of economic tendencies in a certain type of economic system may have little or no validity in connection with a different type of economic system.

Since the student of economics who lives in a capitalistic system is likely to be primarily interested in economic activity as it occurs in this system, this book, like most of the others in the field of economic principles, will be chiefly concerned with describing and explaining economic activity as it takes place under capitalism. But this should not be allowed to give the impression that the forms, methods, and policies of economic activity under this system are inevitable and universal in character, or that all the laws and generalizations to be advanced are equally valid under any type of economic system. Consequently, in addition to describing and explaining the ways in which economic activity is conducted under capitalism, we shall try, from time to time, to suggest how similar matters are handled under other systems, while our laws and generalizations will be examined as to their prospective validity under other economic systems. We shall turn now to a description of the economic institutions and conditions which prevail under capitalism.

INSTITUTIONS OF CAPITALISM

Freedom of Enterprise. According to the dictionary, an institution is (in the present context) a practice, custom, or convention which is a material and persistent element in the life or culture of an organized group. It may or may not be consciously approved by society, and organized and maintained through prescribed rules and agencies. Freedom of enterprise is one of the basic institutions of a capitalistic economic system. It refers to the general right, which each individual has, to engage in any line of economic activity which appears desirable to him. He is not a serf, bound to do the will of his lord and master, nor is it necessary for him to become a farmer or a blacksmith because his father followed the one or the other of these occupations. As far as the government is concerned, the individual is free to move to any part of the country, select any occupation which pleases him, and found and operate a business unit in virtually any field of economic activity.

It is not our function to praise or condemn the institution of freedom of enterprise, but it is significant to note the probable reason for its existence in a capitalistic system. The theory which is ordinarily used to explain and justify the existence of freedom of enterprise is quite simply one of social welfare. That is, the individual, in choosing the field of economic activity in which he will be most productive and useful to himself from the point of view of private gain, will also be selecting the field, it is assumed, in which he will be most productive and useful from the point of view of group welfare. The theory does not assume any high degree of altruism or desire for social service on the part of the individual. He may choose a certain phase of economic activity only because he expects to maximize his own income in that way, but ordinarily such a choice will enable him to make the greatest possible contribution to the production, income, and welfare of society as a whole as well.

Restrictions on Freedom of Enterprise. The social-welfare theory in connection with freedom of enterprise apparently assumes that social welfare will be maximized if we can induce each individual to produce commodities or services which will have the greatest possible value on the market. This assumption, as we shall see in the next chapter, is by no means always justified. Quite apart from this objection, however, capitalistic societies have never been willing to extend complete freedom of enterprise to the individual. That is, it has always been recognized that the individual, in selecting the field of activity which would be most profitable for himself, might well choose something which would be, clearly and by common consent, antisocial. In such cases, governmental units under capitalism have not hesitated to come forward with prohibitions.

Thus the individual's search for a type of activity which would max-

imize his income might lead him to rob or steal, to operate a saloon or a billiard emporium, to pose as a doctor or a lawyer though he lacked adequate professional training, or to sell for his own profit and as his own product a book, a piece of music, or some mechanical device created by someone else, but in such cases the community or group steps in with prohibitions or restrictions. Burglary, robbery, and other predatory activities are forbidden by law. The operation of a saloon or a billiard parlor is licensed for a substantial fee to control the number of such establishments and to ensure that the operators will meet certain minimum conditions and requirements which are imposed in the interests of the public. A doctor or a lawyer must have had a required amount of training and be able to pass state examinations imposed to limit entrance into these professions to persons whose abilities and preparation are adequate. Authors, composers, inventors, and others are protected, by patent and copyright laws, from people who would misappropriate their works or contrivances.

However, restrictions on freedom of enterprise are not limited to activities of a distinctly antisocial type. In the case of certain industries which are considered to be "affected with a public interest" and which operate under severe governmental regulation, such as transportation and public utilities, restriction of freedom of entry is a common development. For example, as part of the federal regulatory system, restriction of freedom of entry has been practiced in the railroad industry since 1920; in motor-vehicle transportation since 1935; in transportation by air since 1938; in transportation by water since 1940; and in the transportation of natural gas by pipe line since 1942. Restriction comes about through the requirement that prospective founders of enterprises must secure permits, or certificates of public convenience and necessity.

After setting up an enterprise in a particular field, the individual may find himself subject to a host of additional regulations which indicate what he should or should not do in carrying on his business. However, such regulations are not really restrictions on freedom of enterprise so long as they only regulate going concerns in the various industries and do not restrict freedom of entry into the industries, though they may often have both effects. In wartime, governmental actions may not only compel firms and industries to give up the production of certain economic goods (such as automobiles) and take up the production of other commodities (such as airplanes) but may prohibit the entry of new firms into certain industries and thus interfere directly with freedom of enterprise. In any case, however, we should note that the restrictions on freedom of enterprise apply ordinarily to all persons who wish to enter the fields of economic activity in question and are not intended to give some persons differential advantages over others. Even with all the restrictions, the individual is much more free than in a system in which his choice of activity is con-

trolled by such matters as race, religion, or social class. The legal restrictions on freedom of enterprise under capitalism are merely exceptions to the rule, and not the rule itself. They are departures from the norm.

Freedom of Enterprise and Economic Opportunity. However, it should be noted that the norm implies only freedom to choose one's occupation or business and freedom of action therein. It contains no guarantee of success in the chosen line of activity, nor does it promise all individuals an equal chance to succeed from the economic point of view. In the economic race the individual must furnish his own equipment. Some individuals enter the contest with the help of adequate training, capital, and business connections, in addition to considerable natural ability. Others have meager ability, and almost no training, capital, or connections. It is a handicap race in which some of the faster runners are placed nearer to the goal while the slower runners must often start from scratch. The economic prizes go to those that reach them first. It should be remembered that, no matter what the politicians say, a capitalistic system offers only freedom of entry into occupations and industries and not genuine equality of opportunity therein.

Some people today question whether the institution of freedom of enterprise means very much to the average individual. To be sure, under it many individuals have advanced themselves from a lowly economic status to one of affluence and power, but success of the self-made variety becomes progressively more difficult to attain as time goes on. As an economic system grows older, as all its valuable land (including natural resources) comes to be appropriated, and as its businesses and industries increase in size and capital requirements, an individual's chance to strike out for himself as an enterpriser and achieve great economic success is thought to become more and more limited. But at any rate the individual may find some consolation in knowing that these conditions also affect other individuals, and he may still prefer freedom of enterprise to any alternative, even though in practice that freedom may allow him only to choose the great corporation for which he will go to work.

Economic Motivation. Economic motivation is a second important institution of a capitalistic system. It means that individuals in this system are usually motivated by the desire for economic or pecuniary gain in their economic activities. They try to behave economically and make the scarce means of production go as far as possible among their various alternative uses in the satisfaction of human wants. Most people do not have to be urged to follow their own self-interest. They are strongly inclined to seek their own welfare, as they conceive it, or that of persons closely connected or associated with themselves, rather than that of other people. And in economic activity the pursuit of self-interest involves trying to get as many economic goods as possible for oneself, or for persons closely identified

with oneself, with little regard for the effects of one's actions on other people.

We are not attempting to picture the individual under capitalism as being completely absorbed in efforts to secure large quantities of economic goods for himself with a minimum of effort or disutility on his part. The famous "economic man," who not only always seeks but succeeds in obtaining maximum income at minimum cost, is a figment of the imagination. Many other factors besides a mere selfish desire for pecuniary gain influence individuals even under capitalism. They are motivated in part by the love of power, by a desire for the admiration and respect of their fellow men in general or of other members of their particular social group, by the pleasure which they find in their work, by the desire to serve their fellow men, or by family affections. For example, we may see a lawyer retire from private practice to accept a position on the Supreme Court, although his salary as a justice may fall far short of the income which he would ordinarily expect to receive as a practicing attorney; or a father select his son for a responsible position in his business when he knows only too well that better economic results might be secured by hiring a stranger for the position; or even a college professor refuse a better-paid position at a distant college because he wishes to remain near his family home. These are exceptions. In general, in a capitalistic system, the desire for economic gain seems to make people work harder and longer than would any other motive which could be substituted for it.

The active defenders of our capitalistic system usually attempt to justify the institution of economic motivation on the basis of social welfare. They assume that the individual, in trying to follow his economic self-interest, will conduct himself in such a way as to confer a benefit upon society as a whole, whether or not he intends to do so. Suppose, for example, a manufacturer wishes to acquire a large income for himself. One way of achieving this end is to excel his competitors in efficiency. In order to lower his costs of production he makes use of the best machines and productive methods available and he eliminates waste at all points in his organization. If his efforts are successful, he can make plenty of money even though he must sell his goods at the price which his competitors charge. In the end, however, his competitors will adopt his superior productive methods, costs of production will be lowered throughout the industry, and consumers will benefit by obtaining the product in question at a price lower than that which formerly prevailed. Thus the manufacturer, in attempting to benefit himself, sets in motion a train of events which produces a benefit for other people, even though this result may have been no part of his intention.

Restrictions on Economic Self-Interest. There are undoubtedly many cases in which such a double benefit results from economic activity on the

part of the individual. In fact, under some conditions, it might well be that this result would follow from most economic activities of the individual, but in our modern complex economic system it frequently happens that individuals, in following their economic self-interest, obtain an advantage for themselves at the expense of the group as a whole. Our manufacturer noted above, instead of increasing the efficiency of his organization, might try to obtain the large income he desires by forming a monopoly and charging high prices, by "sweating" his labor or employing child labor under miserable working conditions for very low wages, or by cheapening his product through adulteration. In such cases the social group is likely to intervene with laws designed to control the growth and functioning of monopolies and trusts, to regulate working conditions and the use of child labor in industry, or to prohibit the sale of impure food and drugs. However, restrictions of these kinds must be regarded as departures from the norm of capitalistic operation.

Private Property. Economic motivation and freedom of enterprise would not in themselves be adequate to insure the operation of a capitalistic economic system if individuals were not protected in the use of the economic goods which they acquired through economic activity. Unless he is assured of being able to control and enjoy the fruits of his labor, why should the individual attempt to follow his economic self-interest and try to select the occupation or business in which greatest quantities of economic goods will be forthcoming? Thus in a capitalistic system there exists the institution of private property. Under its operation the individual is entitled to use and control the economic goods which he acquires, to exclude other people from using them, and to decide usually how they shall be disposed of after his death. The system is likely to include freedom of contract, or the right to bind oneself or one's possessions to the fulfillment of certain future conditions, and contracts freely made between competent and presumably equal persons are enforceable at law. The individual may abuse, as well as use, his wealth if his activities will not interfere with the equal rights of other persons. In every-day usage, the term "private property" is sometimes used to refer to the rights just mentioned and at other times to mean the objects of wealth to which the rights apply, but for our purposes the latter use of the term should be avoided.

Although individuals who own large amounts of wealth often come to revere the institution of private property as one of divine or at the very least of natural origin, it does not seem possible to accept either of these conclusions, for private property is a set of rights conferred by society as a whole upon its individual members. Apparently it is assumed that all of the members of society will be better off if each of them is granted these particular rights. That is, private property is supposed to furnish an incentive to economic activity and to the accumulation of capital and other

durable objects of wealth. In the absence of this institution, economic motivation and freedom of enterprise would mean very little, for economic activity would tend to be limited to the acquisition of commodities for immediate consumption, or of articles which, combining large value and small bulk, could be moved quickly and easily if someone threatened to appropriate them. Since large amounts of capital are essential to the efficient operation of an economic system, the function of private property is an important one. We are not contending that the existence of large aggregations of capital or other wealth in the hands of a few individuals, protected by private property, furnishes an incentive to the individual who is just entering some field of economic activity. Such conditions may be discouraging rather than stimulating, but the institution of private property may at least furnish the individual with an incentive to go and do likewise, for he may be confident that, whatever the extent of his success, he will be protected in its use and enjoyment.

Restrictions on Private Property. While the institution of private property is said to confer on individuals the right of exclusive control and use of the economic goods which they acquire, it is clear that, since property rights are conferred by society as a whole upon its individual members, these rights can be modified and limited by social action. In practice, even under capitalism, property rights are often restricted severely by the action of larger or smaller social groups or governmental units. Thus, in many cities, an individual cannot acquire a piece of land in a residential district and set up and operate a slaughterhouse, factory, or garage upon it. Details of building construction and of location with respect to streets, sidewalks, and adjacent buildings are also frequently regulated. An individual cannot with impunity use his automobile or other objects of wealth in such a way as to damage or destroy the wealth of other people. Moreover, if society decides that too great inequality in the distribution of income is undesirable, the rights of individuals to use and enjoy their money incomes may be limited by a progressive tax applicable to these incomes. In similar fashion, if it is decided that the inheritance of wealth and of claims on wealth is operating to produce an undue concentration of ownership in the hands of a relatively few individuals, this part of the institution of private property may be modified, through inheritance taxation or otherwise, or even eliminated. Outside such restrictions and limitations, a capitalistic society goes to great lengths to protect the possessions of the individual from encroachment by other individuals or by governmental units; hence the restrictions and modifications of private property, like others previously noted, must be regarded as exceptions to the rule in a capitalistic system.

Public and Corporate Property. We should not suppose that, because the institution of private property operates in a capitalistic system, all wealth is owned and controlled by private individuals as such. There is

actually a great deal of public property within most capitalistic systems. Where public property exists, the exclusive control of wealth is exercised by a larger or smaller group of individuals through some governmental unit or other, and the individuals within the group which the governmental unit represents ordinarily have equal rights to use the wealth in question. Streets and roads, sidewalks and street lamps, parks and public buildings are examples to which the concept of public property applies. Moreover, even under capitalism actual productive plants or industries may be owned and operated by various governmental units, as representatives of the body of citizens.

A considerable part of the wealth of a capitalistic economy—some 80 per cent of all business wealth in the United States—is not owned directly either by private individuals or by public bodies, but rather by corporations. From the legal point of view, corporate property is private property, but it has some distinctive features. Private individuals own shares of corporate stock, but these shares are claims only against the general wealth of the corporation and not against specific objects of wealth, and the part owned by one individual cannot be separated from the parts owned by other individuals. Moreover, in most corporations, a relatively small group of individuals manages and controls corporate wealth, while large numbers of stockholders are deprived of that control over their wealth which is ordinarily an important feature of the institution of private property.

Competition. The efforts of individuals to advance their economic self-interest, under the protection of the institutions of private property and freedom of enterprise, result in competition, which is another fundamental characteristic of a capitalistic economic system. Like economic motivation, competition is a way of reacting in economic situations; it is a complex pattern of human behavior. The market for a given economic good is said to be freely competitive only if several general conditions are fulfilled. There must be a large number of both buyers and sellers in the market for the good. The number of buyers and sellers is considered to be large if any one seller can offer more or less of the good, and any one buyer can increase or decrease his purchases of the good, to any extent within his power without affecting the price of the good in the market. There must be no agreements or conspiracies among sellers or among buyers with regard to the production, sale, or price of the good. That is, the buyers and sellers must be really independent individuals or firms. Buyers and sellers must be reasonably intelligent and reasonably well informed concerning market conditions affecting the good in question. Individual buyers and sellers are supposed to be substantially equal in economic strength or bargaining power. Factors of production must be mobile as between industries and occupations and, where possible, between places. Finally, the price of the

good is supposed to be determined by the joint action of the forces of demand and supply without governmental interference.

Clearly, competition in the economic sense is not a natural thing. It is a social pattern contingent upon the existence and operation of the various supporting capitalistic institutions. The active supporters of capitalism usually attempt to justify the existence of competition on the basis of social welfare. Under the operation of competition, efficiency in the operation of industry and business is supposed to be promoted, for economic success will be attained by those enterprises which are efficiently operated, while those which are inefficiently and wastefully operated will be relentlessly eliminated. Thus each enterprise which would survive must make use of the best machines and productive methods available and eliminate waste at all points in its organization, while the removal of inefficient producers from the market tends to leave productive factors in the hands of those enterprisers who can use the factors most effectively. Again, competition is supposed to promote innovation and technological progress. Better productive methods or machines which increase efficiency and lower cost, or superior performance or new qualities of an economic good which enable it to satisfy a human want more effectively than the products of other enterprisers, give certain enterprisers a greatly prized differential advantage over others with respect to income. But such differential advantages are wiped out sooner or later in a competitive industry and the continuing result is the ability of consumers to obtain better and better products at lower and lower prices.

Finally, competition is said to be a kind of regulator or governor of economic activity, a means whereby the productive efforts of numerous individual enterprises are coordinated with the desires of consumers as expressed on the market through prices. Success in competition depends on the ability to give consumers the right amount, quality, and kinds of economic goods. Enterprisers who supply goods which are not suited to consumers' desires, which are of unsatisfactory quality, or which cost more than the similar products of other enterprisers tend to fall by the wayside. If the total output of an economic good is small relative to the effective demand on the market, the existence of a profitable price will furnish the stimulus for a competitive industry to expand production and, if necessary, plant facilities. On the other hand, if competitive producers turn out a total output which is large relative to effective demand on the market, unprofitable operation will tend to force some enterprisers out of the industries and bring about a more suitable adjustment of output to effective demand. The fact that the individual buyer must compete with other buyers in obtaining a certain good will keep him from acquiring it at as low a price as he would like, while competition among sellers makes it impossible for the individual seller to charge as high a

price as he would like to get for the good. Thus competition is supposed to work in conjunction with the other capitalistic institutions to achieve the maximum production at minimum cost of precisely those goods which consumers most desire, in so far as the relative strength of consumers' desires is accurately indicated by the prices which prevail for various economic goods on the market. And, so the advocates of capitalism say, these results mean the maximization of social welfare.

Critics of capitalism are generally unwilling to accept this account of the merits and advantages of competition, for they believe that, as it operates in practice, competition produces several wasteful results and tendencies. Competitive producers wish to be able to handle not only their usual volume of business but also any customers that they may be able to maneuver away from their competitors, and the result, from the point of view of the economy as a whole, is idle productive capacity and duplication of productive facilities in many industries. The desire to obtain a differential advantage over other producers leads enterprisers under competition to produce a wasteful and excessive number of styles, shapes, sizes, and colors of economic goods. Much of our large volume of competitive advertising is considered by these critics to be totally useless from a social point of view, and the competitive exploitation of natural resources is notoriously wasteful and inefficient.

Restrictions on Competition. In some of our markets for economic goods the conditions of free competition are ordinarily closely approximated, as, for example, in the markets for certain agricultural products, such as wheat and corn, and in the securities markets. However, in the markets for many other economic goods, competition is present only to a limited extent if at all. The governments of capitalistic economies place many restrictions and limitations on economic activity, and interferences with competition are the result. Consumers are not left to depend entirely on the tender mercies of competition to provide them with foods and drugs of adequate purity and quality, for there are laws which set up certain standards in these matters. Again, laws which establish minimum wages and maximum hours for workers in many industries may seriously alter the results which the ordinary operation of the labor market would produce. In some cases, governments may actually prohibit the operation of competition in certain fields of economic activity, as, for example, when a governmental unit grants to a public-utility enterprise the exclusive right to sell a certain commodity or service in a given market area. Governmental activities which interfere with competition usually modify the rights involved in freedom of enterprise and private property as well; hence we see that governmental prohibitions and restrictions ordinarily affect the operation of several institutions at once instead of affecting only a single institution.

those which result from the voluntary actions of individuals or groups of individuals. Groups of workers in various labor markets may organize unions which impose severe restrictions on membership in the hope of limiting competition and exacting a monopoly price for the services of the members. On the other side of the picture are the many capitalistic industries owned and operated by a relatively few large firms, instead of the many small firms which strict competitive conditions require. In some cases, a single large firm will control all or a large part of the productive capacity of an industry or a few great firms will act in close harmony to achieve monopolistic control. Even when the few large firms in an industry do not enter into any formal agreement or combination, they usually cannot be as independent of each other as competitive conditions would necessitate. In addition to conditions of outright monopoly and quasi monopoly, many industries operate under conditions of monopolistic competition. The firms in these industries turn out products of the same general family or type, but not identical goods. They "differentiate" their products on the basis of mechanical gadgets and features, brands, packages, and other devices, and attempt to convince customers of the superiority of their particular products through extensive advertising and salesmanship.

Firms which operate under conditions of monopolistic competition or quasi monopoly, unlike the individual sellers under competitive conditions, cannot increase or decrease production on the assumption that these actions will not affect the price which they can get for their product, and so they do not produce a given amount of their good and then throw it on the market for what it will bring. Instead, they are inclined to decide in advance what prices they will charge and to vary production according to the possibilities of sales at these prices. Price competition between the firms is not very likely to occur, since a cut in price by one producer is likely to be matched quickly by similar reductions on the part of other producers, and no one gains by cutting prices. Under monopolistic competition, one cigarette is said to be easier on the throat because "it's toasted," another is especially soothing to tired nerves, and another lacks a certain irritating ingredient common to other cigarettes, but the various companies charge the same price for cigarettes of the same general quality. One gasoline was said to enable the car owner to get the most out of his wartime rations, while "ground crew" service was furnished along with high-quality gasoline by another company, but no large oil company offered to sell a standard grade of gasoline at a price significantly lower than that which was charged by its competitors.

Though comparatively few of the industries of the United States operate under strict competitive conditions, we should not jump to the conclusion that our economy is not competitive at all. Monopolistic competitors in an industry, besides competing with each other on the basis of quality and

product differentiation, must compete for customers with the enterprisers of industries which produce other families of products intended for the satisfaction of the same general want on the part of consumers. For example, the monopolistic competitors in the manufacture of electric refrigerators compete with those who produce the old-style ice refrigerators. They also compete with enterprisers in completely unrelated industries for the limited money incomes of consumers in general. Many individual consumers may want to purchase both new electric refrigerators and new radios. If they have enough money to make the one or the other purchase, but not both, the sellers of electric refrigerators are seen to be competing in this case with the sellers of radios.

Monopolistic competitors in any industry compete with each other, and with enterprisers in all other industries, for supplies of the scarce agents of production. The firms in any given industry also compete on the basis of technology. That is, they continually try to develop improved machines and methods of production which will both lower their own costs of production and render obsolete the machines and methods of their competitors. Competition of this type has been of increasing significance in recent years. Even monopolists, who do not have to face competition within their own industries, must compete with enterprises in other industries for the limited money incomes of consumers and for supplies of the scarce agents of production. Finally, some of our industries operate under conditions which approximate competition in the strict sense of the term. Thus our modern capitalistic system is said to be competitive in the sense that it is more nearly competitive than anything else. Just as the existence of economic motivation in a capitalistic economy does not mean that all individuals are always and completely motivated by the desire for economic gain, so the existence of competition in the same system does not require the existence of *perfect* competition in all fields of economic activity.

Cooperation. While the operation of a capitalistic economic system has been described as fundamentally competitive in character, it is nevertheless true that it contains a considerable element of cooperation. Some of this cooperation is of the conscious variety. Although the ice dealers in a given market area may compete with each other for customers, they may at the same time cooperate through their trade association in advertising intended to convince consumers of the advantages of ice-cooled refrigerators over mechanical refrigerators using electricity or gas. In similar fashion, individual candy manufacturers may compete to sell the consumer chocolates made by Whitman, Schrafft's, or Page and Shaw, but they may also cooperate in a campaign to educate the consumer as to the desirability of candy eating in general so that he will consume more of their luscious products than formerly. Not only do business firms cooperate

with each other for certain purposes but the firms themselves may be organized on a cooperative basis, as we shall see in Chapter VIII.

However, the most important kind of cooperation in our economic system is that which is largely unconscious in character. Unconscious cooperation is merely another name for specialization, or the division of labor. How is it possible, we may ask, for the college professor to spend his working days talking in the classroom, marking papers, attending committee meetings, and engaging in research, and yet satisfy as many of his wants as he does? It is because producers in many other industries cooperate to furnish him with food, clothing, shelter, transportation, recreation, and many other commodities and services. However much or little competition there may be within the many lines of production, they work together toward the final satisfaction of human wants. Thus the various phases of economic activity are closely interrelated and interdependent in our economic system.

INSTITUTIONS OF SOCIALISM

Private and Public Property. The comparisons, which will be made later on, of the methods and principles of economic activity under capitalism with those which are significant under other types of economic systems, will be more meaningful if we pause briefly at this point to suggest the institutions and conditions under which economic activity would be or is carried on in these other systems. Turning first to socialism, we may note that socialism has been described as "an economic organization of society in which the material means of production are owned by the whole community and operated by organs representative of and responsible to the community according to a general economic plan, all members of the community being entitled to benefit from the results of such socialized planned production on the basis of equal rights."[1] This statement apparently suggests that, under socialism, the rights of private property would be limited to consumption goods, since productive wealth (land and capital) would in general be owned by society as a whole.

However, some socialists say that the social ownership of the means of production would be limited to the land and capital used in "large-scale production." That is, aggregations of land and capital which were so large as to require the use of hired labor in their operation would have to be socially owned, but small amounts of land and capital, which could be operated in production by the owner (and perhaps other members of his immediate family) could perhaps be safely left in the hands of private owners. Sometimes it is even contended that certain industries, which operate quite well under private ownership and which are not well suited to governmental ownership and operation, might well be left in the hands

[1] From *Economics of Socialism*, by H. D. Dickinson. Oxford, Clarendon Press, pp. 10–11.

of individual owners or cooperative groups. These modifications would apparently permit private individuals even under socialism to own and operate small farms, stores and repair shops, and shops for handicraft production, and it is possible that certain fields of economic activity, such as agriculture and retail merchandising, might be carried on entirely by private owners or cooperatives. Nevertheless, the general conclusion is that the great bulk of land and capital would be owned by society as a whole under socialism.

In the socialistic economy of Soviet Russia, the ownership of the material means of production has been taken almost completely out of the hands of private individuals. By the end of the Second Five-Year Plan (1937), some 98.7 per cent of the land and capital of the system had been brought into socialized or collective ownership.[2] The land of the country was nationalized as of February, 1918, and is completely owned by society as a whole, while practically all capital goods in industry, and the heavy capital goods in other fields of activity, are similarly owned. The numerous cooperative or collective farms of the country own some farm buildings, draft animals and other livestock, some machinery, and stocks of seed, and have the right of "perpetual use" of their land, though tractors, combines, and other types of heavy agricultural machinery are publicly owned and are furnished to the collective farms on a kind of rental basis. The fairly numerous marketing cooperatives and producers' cooperatives in handicraft production own certain limited quantities of capital goods, and very small amounts of capital are in the hands of individual peasant farmers and handicraft producers.

Freedom of Enterprise. Under ideal socialism, the people as a whole, through various units of government, would operate as well as own most of the productive wealth of the country. This does not mean that the *central* government of the system would own and operate all industries, for many types of production might be allocated to other governmental units and some cooperative and private enterprises might be allowed to exist. However, since the private and cooperative enterprises would be expected to function in conformity with the general economic plans of the system, it may be said that freedom of enterprise would exist, for all practical purposes, only for the people as a whole through their government. On the other hand, the individual citizens of the system would continue to enjoy considerable freedom of choice in selecting their occupations under socialism.

This description is substantially applicable to the socialistic economy of Soviet Russia. Most important types of economic activity are operated, as well as owned, by the federal, republic, district, or local governments. However, in agriculture the roughly 250,000 collective or cooperative farms cul-

[2] M. T. Florinsky, *Toward an Understanding of the U.S.S.R.* New York: The Macmillan Company, 1939, p. 173.

tivate over 90 per cent of all the land under cultivation, and in 1938 included some 18.8 million peasant households, or 93.5 per cent of the total.[3] Some 3 million persons are members of various other cooperative enterprises in handicraft and service production, and many thousands of consumer cooperative organizations operate retail stores and are served by cooperative wholesaling organizations. In addition, there are numerous but rather unimportant private enterprises in farming, handicraft production, and service trades. However, the cooperative and private enterprises in Soviet Russia do not really enjoy freedom of enterprise in the capitalistic sense, for their activities are covered by the economic plans of the country and their operation is quite rigidly controlled by means of prices, taxes, interest rates, rentals, and other factors which are directly under governmental jurisdiction. Freedom of enterprise really exists only for society as a whole, though Russian workers ordinarily have a large degree of freedom in choosing their occupations.

Economic Motivation. An ideal socialistic economy would rely on economic motivation to some extent, but the significance of this institution would be much less than under capitalism. The profit motive itself would be virtually eliminated, since private individuals, with limited exceptions, would not be allowed to own and operate enterprises for private gain. Since productive wealth would be owned very largely by governmental units, as representatives of the people, individuals would not be motivated by a desire to accumulate land and capital in order to receive rent or interest as private income. Most individuals would work for some governmental unit or other and would receive their money incomes in the form of wages. This is where economic motivation comes in, for most socialists would permit moderate differentials in wages as between industries and occupations, though there are a few socialists who hold out for a strictly equal distribution of income. By moderate differentials in wages, socialists probably mean that the highest wage paid to anyone would not be more than ten to fifteen times the lowest wage paid to anyone. Such differentials are large in comparison with perfect equality, but are very modest in comparison with the differences in individual incomes which prevail under capitalism.

Though economic motivation would be retained to some extent under socialism, such a society would try to emphasize other types of motivation. Individuals would be supposed to work for the good of society as a whole, for prestige (or the esteem and admiration of their fellow men), or for positions of power in the economic system. Under capitalism, prestige and power are often sought indirectly through obtaining a large income and accumulating wealth, but under socialism they would be separated rather thoroughly from the seeking of economic gain. Public honors would be granted for unusual accomplishments in production, and security against

[3] *Ibid.*, p. 200.

unemployment, disability, and other economic misfortunes might serve as a spur to productive effort. Underlying everything else, of course, would be the public power of compulsion, with individuals liable to be penalized for unsatisfactory work, and the ultimate requirement that every able person must toil if he wishes to eat.

In these matters, Soviet Russia lives up to the socialistic model at least in a general way. While the profit motive as such is almost completely eliminated, economic motivation is relied upon to a considerable extent in Soviet Russia and there are fair-sized differentials in money wages. Just how large the differentials in money income are is a matter of some dispute. They are sometimes pictured as being as small as 18 or 20 to 1 from top to bottom, while in other cases the claim is made that they run as large as 60 or 70 to 1. In either case, the differentials are much smaller than those which ordinarily prevail under capitalism. Quite apart from their money incomes, the individuals of the Russian system receive part of their real income from the government in the form of direct grants of commodities and services, and these grants may increase, decrease, or merely maintain the differentials which exist on the basis of money incomes, according to the manner in which the grants are distributed among the workers with different levels of money income.

The Soviet Russian economy also relies to some extent on other motives, such as altruism, enthusiasm for socialism and economic planning, and socialist competition, or the efforts of the workers in various socialized enterprises to outdo each other in the matter of increasing production. A considerable range of public honors is available for conspicuous achievements in production, and some of these honors carry certain economic rewards with them as well. Some economic positions afford much more prestige and power than others, and give security against unemployment, old age, and other economic ills. Many penalties are also provided for workers whose attitudes, efforts, or results are particularly unsatisfactory, and the public power of compulsion can be used as a last resort to get the workers to function in the manner and in the places desired by the economic planners.

Competition. A great reduction in the importance of competition in economic life is one of the features of an ideal socialistic system. Individual workers would presumably compete to get into the better-paying and more pleasant occupations of the system, and consumers would compete for the available limited supplies of various consumers' goods and services, since the socialistic intention is to give consumers considerable freedom of choice in spending their money incomes, but that is all. In other respects, the governing influence of competition would be replaced by the dictates of economic planning. This is the situation which exists in Soviet Russia if we add to the types of competition just mentioned the "socialist competition" to which

we referred in another connection and some unauthorized competition of enterprises and industries to secure supplies of materials, fixed capital goods, or labor. A socialistic economy is in general intended to be a single great cooperative enterprise.

INSTITUTIONS OF COMMUNISM

If we sought to understand the differences between socialism and communism by examining the platforms of the Socialist and Communist parties in the United States, we might reach the conclusion that these differences are not very striking and significant. However, there are important differences between socialism and communism as ideal theoretical systems. In the first place, communism would go further than socialism in modifying the capitalistic institution of private property. Under communism, consumers' goods as well as land and capital would be owned by society as a whole. Of course, various consumers' goods, such as groceries, clothes, and toothbrushes, must be allowed to pass into private possession in order that consumption may take place, but the basic title even to such things would remain vested in the entire social group. Freedom of enterprise, as a right of individuals, would be completely eliminated under communism. Individuals would presumably remain free to choose their own occupations, but these choices would be made entirely on some other basis than the desire to receive a large money or real income.

Thus in an ideal communistic system, economic motivation would be completely eliminated. Each individual would choose that occupation in which he could be most useful to society as a whole, would produce to the best of his ability, and would receive real income according to his needs, presumably by helping himself to the various economic goods which would be available in large quantities in public storehouses. While the distribution of income according to needs would not result in a precisely equal distribution of real income among persons, unless we assume that all persons have equal needs, economic inequality between persons would be almost entirely eliminated under communism and any differences in income which existed would not be based on productiveness or choice of occupation. If there were no differential wages for which workers could compete, and if all consumers' goods and services were so plentiful that each person could have all he needed, then clearly the last vestiges of competition as a capitalistic institution would be removed under ideal communism. Complete social cooperation would be the order of the day. From this brief discussion, we can see why the institutions of Soviet Russia have been described as socialistic in character rather than communistic, even though Russia is widely known as a communistic system and its operation as an economic system is well-nigh completely in the control of the so-called Communist party.

INSTITUTIONS OF

In the fascist economies of Italy and Germany, the national leaders held that the goal of their system was the greater glory of the state, or of the nation as an entity separate from the numerous individuals who composed it at any particular time, rather than the economic welfare of the citizens as individuals. And national glory or prestige was to be secured by means of wars of aggression. Economic policies were followed or abandoned not for their own sake or because of their effect on individual welfare but according to whether they appeared to be consistent or inconsistent with the general goal of the systems. Such systems did not need to make any changes in the *nominal* character of economic institutions as such. They needed merely, by intervention and restriction, to control the operation of these institutions so that their effect would be to further the aims deemed appropriate by the leaders.

Thus the differences between capitalism and fascism did not lie in the basic institutions of the systems, for these institutions were nominally the same in both systems. Under fascism we saw no such complete concentration of the ownership of productive wealth in public hands as is presupposed by a system of socialism or communism. The various governmental units under fascism did own and operate some industries, but the great bulk of productive wealth was left in private hands. However, the private owners of productive wealth were not supposed to think that they had any sacred property rights, for they held their wealth only on the sufferance of the "leader," Herr Hitler or Signor Mussolini. The uses which could be made of productive wealth were closely controlled, and private wealth was even expropriated on occasion. Thus when the government or party became short of funds, examiners could be sent to go over the books of private concerns for many years past, and heavy fines could be levied for any false entries or mistakes in bookkeeping, however trivial and insignificant. There were no laws or courts capable of preventing such arbitrary governmental actions, or of interfering with the capital levies which were placed on landowners or enterprisers on other occasions.

On paper, the fascist systems made much of individual initiative, economic motivation, and freedom of enterprise as devices for securing the efficient operation of their economies. Unless and until he was interfered with, the individual was free to be a business enterpriser, produce any commodity or service which he liked in any quantity which seemed most appropriate, hire and fire labor, secure supplies of other productive factors on as good terms as he could get, make as much money as he could, and spend or save his income at his pleasure. Under these conditions, it would seem that the operation of the economic system would have been highly competitive in character, and that private businessmen would have been able

to make their various economic decisions on the basis of price movements and relationships.

Actually, interference with the private operation of business and industry existed to an almost unbelievable extent under fascism. The government would restrict, or even prohibit altogether, the entry of individuals into certain industries and businesses. The government controlled the prices at which commodities and services could sell or at which labor or other productive factors could be obtained, it regulated the marketing of economic goods, it set up production quotas or hours of operation for plants and enterprises, it rationed raw materials and supplies, it compelled employers to hire and fire labor along party lines without regard for ability, and it controlled the use of foreign exchange, the importation of all sorts of economic goods, and the exportation of finished products. Under these conditions, business enterprises could seek profits but, if profits were made, the government decided whether they could be paid out to the owners of the businesses and to what extent, and it could require the enterprises to invest their earnings in the securities of new plants that were being set up for purposes of economic self-sufficiency, or in government bonds. Differentials in wages existed between the various occupations and industries, and the workers sometimes had considerable freedom in their choice of occupations, but in the later years of fascism there appeared a tendency to freeze the workers in their jobs and to mold the class structure of society into rigid and semipermanent castes or groups. In any case, in view of the great levies which the fascist governments made against the wages of workers, it is questionable whether there remained any great degree of economic motivation for them.

Fascism, then, appeared to be different from capitalism in that an all-powerful central government, unrestrained by constitutional or other limitations, interfered with and controlled economic and other activities to an enormous degree for the purpose of directing them toward the achievement of whatever goals seemed desirable to the leaders of the party and the government. Fascism had the institutions of capitalism in name only, for it refused to let them operate in ordinary capitalistic fashion lest they enhance the welfare of the citizens as individuals instead of making for the different goal or goals which were attributed to the semimystical state or nation.

QUESTIONS AND PROBLEMS

1. What is freedom of enterprise and why is it relied upon as a basic institution in a capitalistic economic system? How and why is freedom of enterprise restricted in such a system?

2. Distinguish between freedom of enterprise and equality of opportunity.

3. What is the connection between economic motivation and freedom of enterprise? Explain.

4. "Under capitalism, the desire for economic gain makes people work harder and longer than would any other motive which could be substituted for it." Discuss.

5. "In operating on the basis of economic motivation, individuals usually benefit society as a whole as well as themselves." Do you agree? Explain.

6. "Freedom of enterprise and economic motivation would mean very little if the institution of private property did not exist under capitalism." Explain.

7. "Since the right of private property is granted by society as a whole to its individual members, it is one which society can modify or limit if it appears necessary to do so in the public interest." Explain and illustrate.

8. How does private property differ from public and corporate property?

9. Outline the several important functions which the institution of competition is supposed to perform in a capitalistic economy.

10. "Even in our capitalistic economic system, competition is interfered with and limited both by governmental activities and those of private individuals." Explain.

11. "While competition is limited or interfered with in many fields of economic activity, it is still correct to consider our economic system as fundamentally competitive." Do you agree? Explain.

12. Distinguish between conscious and unconscious cooperation.

13. Indicate briefly the probable fate of each economic institution of capitalism in an economy of modern socialism.

14. "The economic institutions of capitalism would be completely eliminated in a communistic system." Explain.

15. "The economic institutions of the fascist countries were nominally the same as those of capitalism, but in practice the functioning of these institutions was severely modified." Explain.

REFERENCES FOR FURTHER READING

Blodgett, R. H., *Comparative Economic Systems.* New York: The Macmillan Company, 1944, Chap. 2.

Dixon, R. A., *Economic Institutions and Cultural Change.* New York: McGraw-Hill Book Company, Inc., 1941, Parts 1 and 3.

Loucks, W. N., and Hoot, J. W., *Comparative Economic Systems.* New York: Harper & Brothers, 2nd ed., 1943, Chaps. 2 and 13.

Marshall, Alfred, *Principles of Economics.* London: Macmillan & Company, Ltd., 8th ed., 1920, Book I, Chap. 1.

IV

The Price System

The economic institutions of capitalism, such as private property, freedom of enterprise, competition, and economic motivation, provide the framework within which economic activity takes place in such a system, but they do not specifically determine the nature and content of this economic activity. They indicate what the individual may do, rather than what he shall or must do. That is, if an individual is just starting out in an economic activity, the existence of these institutions does not determine for him what economic goods he will produce and in what quantities, what quantities of productive factors he will use and the proportions in which he will combine them, who will purchase his products and at what prices, or whether he should best spend or save certain portions of the income which he receives.

However, under capitalism and in all other types of economic systems, certain basic decisions must be made. In every economic system it is necessary to determine, in some fashion or other, (1) the kinds and quantities of various economic goods which are to be produced; (2) the way in which the various agents of production are to be distributed among the many industries and fields of productive activity; (3) the way in which the currently available stocks of consumers' goods are to be distributed among consumers; and (4) the total quantity of each reproducible agent of production which is to be available in the long run. These matters are determined in different ways in different economic systems. Under capitalism they depend very largely upon the reactions of individuals and enterprises to stimuli provided by changing price relationships. Thus we should add the price system to the list of institutions which are fundamental to the operation of a capitalistic economy.

THE FUNCTIONS OF THE PRICE SYSTEM UNDER CAPITALISM

The Control of Production. In the first place, price relationships control the kinds and quantities of economic goods which are produced in a capitalistic system. If consumers desire an economic good, such as radios,

very strongly and the quantities being produced and sold at the time are small relative to the public demand, it is to be expected that the prices of radios will be temporarily well above their production costs per unit. Such a favorable relationship of prices to costs will stimulate the manufacturers of radios to produce more of them and may tempt the enterprisers to increase their plant and equipment so that they can continue to expand production. At the same time new enterprisers are likely to be attracted to the radio industry. Moreover, the favorable relationship of prices to costs will permit these enterprisers, both old and new, to bid effectively for land, labor, and capital and to attract and utilize necessary additional amounts of these agents in the production of radios. In the long run, under competition, the production of radios (or other economic goods) will thus eventually be adjusted to the relative strength of the demand for the goods on the part of consumers.

The same forces produce the opposite result if more bicycles, let us say, are produced than consumers desire to buy at the going price. In such a case the price of bicycles may well fall below their production cost per unit. Such an unfavorable relationship of price to cost will usually cause enterprisers to limit production and will make it difficult in the long run for them to attract to their industry additional quantities of the agents of production or even retain what they already have. Moreover, some enterprisers may be driven out of the industry. Eventually under competitive conditions the productive capacity and output of the entire bicycle industry will be cut down and adjusted to the relative strength of the demand for bicycles.

Thus a capitalistic system has no dictator or all-powerful planning agency to specify which economic goods may be produced and which may not, or to determine the relative quantities of various economic goods which should be produced. On the whole the individual is left free to produce any goods which he selects and in any quantity which seems appropriate to him, while the total decisions as to the kinds and quantities of economic goods to be produced are merely the aggregates of the decisions of individual enterprisers—decisions made on the basis of price relationships. And the net result is supposed to be that consumers will receive just the kinds and quantities of economic goods which they desire.

The Allocation of the Agents of Production. The second function which is performed on the basis of the price system under capitalism is clearly implied in the control of production, for the allocation of the available quantities of the scarce agents of production among firms and industries also depends upon price relationships. Suppose a piece of land in the downtown area of a city is vacated and a tenant for it is desired. What will determine the use to which the land will be put? Will it be used as a parking space, or as a site for a theater, or a department store, or a hotel?

Other things being equal, the owner of the land will probably turn it over to the enterprise or industry able to pay the largest amount for its use. In the long run, the entire land supply thus tends to be distributed on the basis of the sums which various enterprises and industries are able and willing to pay for it or for its use. No conscious attempt is usually made to see that the available quantity of land is awarded to those firms or industries whose operation is most necessary to the "social welfare." Rather, it is generally assumed that those industries which can pay most for the land are those which should get it from the social point of view.

When savers make funds available for capital investment, the question of whether these funds will go into capital goods for the production of textiles, automobiles, furniture, plumbing fixtures, or something else is decided in most cases on the basis of the amounts of interest which the various fields of production are able to offer for the use of the funds. There may be some hardy souls who, in spite of attractive interest rates, will refuse to entrust their savings to, say, distillery operators because they think that the production of intoxicating beverages should not be encouraged. Ordinarily, however, those who have savings to invest are quite indifferent to the nature of the products turned out by industries which bid for savings, and in lending out their savings they are only concerned with obtaining the highest possible rate of interest consistent with safety. Even capital goods themselves, to the extent that they are mobile and capable of being used in different types of productive activity, will ordinarily be turned over to the enterprises which can offer most for their use.

Finally, it should be said that the available units of labor also tend to distribute themselves at least in part on the basis of the prices which are obtainable for their services in various occupations and industries. This does not mean that the worker, in considering the desirability of shifting from one occupation to another, will be influenced exclusively by the money wages which are obtainable. He will also give attention to such matters as the pleasantness or unpleasantness of the occupations, the risk or danger involved in each, the relative prospects of advancement, and the probable steadiness of employment. However, workers do shift from one occupation to another, and the principal reason for these changes, though not the only reason, is the desire for high rather than low wages. Many industries which now employ large numbers of workers, such as those producing automobiles, radios, airplanes, and air-conditioning equipment, did not even exist years ago, and they have been able to attract their labor supplies largely because they could offer workers higher wages and better conditions of employment than other industries. The many plants which were constructed for the production of armaments and war materials during World War II obtained their necessary supplies of labor in similar fashion.

The Apportionment of Consumers' Goods. Economic goods command prices because they are scarce in relation to the desires of people for them. Since there are never enough consumers' goods to satisfy all wants or to permit consumers to help themselves freely to the goods, some method or other must be found of apportioning the available supplies of consumers' goods among the people who want them or of holding down the buyers to a number who can satisfy their wants from the existing stocks of goods. Under capitalism the rationing of consumers' goods among the persons who desire them is a third function based on the price system. If the available stock of apples is unusually small, a continuation of the old price would cause the sellers to be swamped with prospective buyers, but if the price is raised sufficiently, the number of buyers will be restricted so that the market demand for and the market supply of apples will be in equilibrium once more. If the available stock of wheat is unusually large in relation to the market demand, the continued charging of a high price would result in the piling up of large quantities of unsold wheat. However, if the price is lowered sufficiently, the number of buyers who come forward may be increased until the quantity of wheat which buyers are willing to purchase is exactly equal to the quantity which sellers desire to sell.

It should be emphasized that rising and falling prices of consumers' goods serve to *ration* the limited available quantities of these goods among consumers. Every rise in the price of a consumers' good limits the number of people who have access to that product, and every fall in the price of an article increases the number of people who may enjoy it. Thus we ration consumers' goods ordinarily by raising or lowering prices and leaving people free to buy as many or as few units as they can at these prices. In times of emergency, however, the necessary rationing can be accomplished by another method. Prices can be kept relatively constant while the quantities which the various consumers are allowed to purchase are increased or decreased as the available quantities of the goods become larger or smaller. This rationing by means of physical quantities of goods might be just as efficient as rationing by means of price changes, but it would not accomplish exactly the same results. When rationing is carried on by means of price changes, the available goods go to the people with the necessary purchasing power to buy them, which means that those who have little purchasing power may not get any of the goods at all. On the other hand, when prices are kept stable and consumers are limited as to the quantities which they may buy at these prices, it is possible to make certain that all the prospective buyers can get at least small quantities of the goods which they desire.

The methods also differ in that rationing through quantities often gives the appearance of greater scarcity than rationing through price changes. When the price of a good is kept at a low level in spite of the fact that

the product is very scarce, the appearance of a quantity of it in the market may induce many buyers to stand in long lines waiting for hours in order to purchase the small amounts which they are allotted. When rationing is accomplished through price changes, the fact that the stock of a good is limited will not cause people to stand in line to purchase it if its price goes up sufficiently. Buyers with inadequate purchasing power will know that they cannot purchase the good in question and will realize that there is no point in standing in line outside the stores in which the good is available for those who can afford it. Thus rationing by means of price changes makes it appear that any quantity of a good, however limited, is adequate to care for the desires of the people who wish to purchase it, though in reality many people who want the good are excluded altogether from its consumption.

Controlling Total Amounts of Productive Agents. We have already seen that under capitalism price relationships control the apportionment of existing amounts of productive agents among the various industries and fields of production. It is also true that price relationships control, where possible, the total amounts of agents of production in existence. That is, a high rate of payment for an agent of production may lead to an increase in its quantity, while a low rate of payment may lead to a decrease in its quantity.

However, it is not always possible to create more of an agent of production. As prices and rents go up, some soil and other natural things which were formerly too poor to be worth using may become a part of the effective supply of land, but there will be no increase in the quantity of natural resources in existence. In the case of labor, the total number of workers available for production may increase or decrease over long periods of time, but it is difficult to estimate the extent to which the number of workers is dependent in the long run upon the wages paid for labor.[1] Some economists of many years ago thought that the total number of workers responded rather directly to wage rates. They professed to see a natural tendency on the part of the wages of labor to be at a level that would enable the workers to live, be physically able to work, and raise large enough families to replace themselves in the laboring population. If for some reason wages went above the so-called subsistence level, workers would have more children than formerly, or more of the children who were born would be able to survive and eventually become a part of the laboring population. Hence, as the result of the increased numbers of workers, wages would

[1] In this section we are dealing with the *long-run* relationships between rates of remuneration and quantities of productive agents. In the short run it is quite possible that increases and decreases in wage rates may affect the number of persons offering themselves for employment, the hours per day or week that people are willing to work, and the intensity of effort put forth by the workers. However, in the long run, the most important factor affecting labor supply is found in the growth or decline of numbers, or population.

fall back to the old level or below. If wages went below the subsistence level, workers would have smaller families, or at least fewer children than formerly would live long enough to become additions to the labor supply, and wages would again become sufficient to provide subsistence.

If these tendencies actually existed, the payment of high wages would eventually result in an increase in the total number of workers available for productive activity, while low wage rates would eventually bring about a decline in the total number of workers. However, for reasons which will be explained later, hardly anyone today has any confidence in this theory of the relationship between wage rates and labor supply, and some writers even put forth a theory which is almost directly opposed to it. For the present it will be sufficient to say that increases in population and the labor supply occur because of a great number of factors—not all of which are economic in character by any means—and that it is virtually impossible to single out one factor, such as wage rates, and estimate its effect on population and total labor supply.

Capital seems to be the factor of production whose total quantity is most responsive to changes in price relationships. Of course the savings which are necessary for capital formation depend upon a number of factors besides the desire to receive interest. Saving occurs, for example, because individuals wish to provide for their own old age or for their heirs, because they want to keep a reserve on hand against the various financial emergencies which may come up suddenly, because they wish to acquire power or prestige, or because they have such large money incomes that they are unable to contrive ways of spending their entire incomes on their personal needs and desires. These factors indicate only that some saving would occur even if interest were not paid; they do not show that the total quantity of savings would remain unchanged if interest were abolished. There are many units of individual income which are just about on the margin of being saved or not being saved. An increase in the prevailing rate of interest will ordinarily cause the owners of such marginal units to forgo present consumption and save larger amounts out of their incomes, while a decrease in the prevailing rate of interest will result in the spending of more of these units of income for current consumption and the saving of fewer units than formerly.

Thus not all units of savings need be directly dependent upon the size of the prevailing interest rate in order that a change in this interest rate may bring about a change in the total quantity of saving and capital formation. The chances are that an increase in the interest rate will bring about sooner or later an increase in the total quantity of capital funds at the disposal of society, while a decrease in the interest rate will eventually cause a decrease in the total quantity of capital funds available for use in production. The magnitude of the total supply of capital funds, as well as the

distribution of the existing quantity among industries, tends to depend upon price relationships in the long run.

OBSTACLES TO EQUILIBRIUM

If all the tendencies which we have been describing worked themselves out perfectly under capitalism, the result would be complete economic equilibrium on the basis of price relationships. In practice, however, there are many obstacles to the attainment of full equilibrium. The preceding description of the operation of the economy on the basis of prices assumed, explicitly or implicitly, the existence of certain conditions which are only roughly approximated in the operation of an actual capitalistic economy. These assumptions include competitive relationships among buyers and sellers of goods, the complete mobility of the productive agents as between places and industries, and the absence of governmental interference with the processes of the market. When these conditions are imperfectly realized, the results produced by the operation of the actual price system do not conform closely to those of the theoretical model.

Immobility of the Productive Agents. In actual practice, the agents of production are never completely mobile. Once land has been committed to a particular enterprise or industry, it is ordinarily tied up for a considerable period of time and cannot be shifted quickly to other enterprises or industries even though more profitable opportunities for its employment arise. In similar fashion, when capital funds have actually been invested in fixed plant and equipment in some industry or other, and these capital goods will last for a number of years, a change in price relationships which would make it possible for these funds to earn more interest in some other type of investment will not cause the funds to move at once to another industry. Hence, in the shorter periods of time, a large part of the economy's capital funds is likely to be immobile, and we cannot expect to find the distribution of the funds among industries to be perfectly suited to the price relationships which exist at a given time. However, it remains true that, in the long run, all capital funds are mobile and none are tied up in fixed capital goods. If savings will command more net interest in some uses of funds than in others, or in some places than in others, after allowing for risk and costs of administration, it is to be expected that savings will move in the long run from occupations or places where net interest is low to occupations or places where net interest is high. That is, as capital goods wear out and create funds for their replacement, these funds will be transferred to other occupations or places instead of being reinvested in new capital goods of the original type.

It must be admitted also that price relationships operate less effectively in distributing workers among industries and places than in distributing land or capital funds. The fact that the wages of coal miners are low in

relation to those of radio announcers will not cause the miners to come out of the ground in a body and mob the radio stations; nor will the fact that doctors receive higher incomes than bricklayers cause many bricklayers to lay aside trowel and mortar for the stethoscope and scalpel. The simple fact is that the labor supply does not consist of a single large mass of like units of labor, all of which engage in competition for the entire range of available jobs. Instead, because the labor supply is stratified into a number of groups or grades, the members of one group find it most difficult, if not impossible, to compete for the jobs which are open to the members of a higher group. The nature of these labor groups, as well as the reasons why movement between them in an upward direction is very difficult, will be discussed later in connection with the theory of wages. For the present, their existence, and the extent to which they interfere with the distribution of workers among industries on a price (wages) basis, must be taken for granted.

Even if we limit the discussion to the various occupations which are open to the members of a single labor group or grade, we shall find that workers do not always change occupations in response to changing wage rates. In some cases, of course, workers may continue to accept a certain wage, even though a higher one is available in some other occupation, because they are not aware of the various opportunities which exist for their employment or because their movement is prevented by powerful unions. In other cases, they may know of the higher wage rates to be obtained in other occupations and may be qualified for these other jobs, only to be unable or unwilling to take them because the better-paying jobs are out of commuting distance. The workers may be unable to stand the expense of moving themselves and their families to another place, they may have formed an attachment for their present location, they may have children in school whose opportunities for education would suffer in the new community, or they may have homes which they could not dispose of without incurring a considerable loss.

Thus the distribution of the agents of production among industries and places on the basis of prices is not an accomplished fact under capitalism but merely a tendency which operates slowly and haltingly through time. The relative immobility of the agents of production obviously has an important effect upon the success with which production can be controlled on the basis of price relationships. Favorable or unfavorable price-cost relationships do afford a stimulus to the expansion or contraction of the production of particular economic goods, but adjustments of production to changes in demand cannot occur quickly and smoothly because of the immobility of the factors of production. And, all too often, before an adjustment of production to one change in demand can be completed, another

change in demand will occur which will render the attempted adjustment of production inappropriate.

Governmental Interference. Governmental interference with and control of economic activity may alter considerably the total productive and other results achieved by a capitalistic economy. Even in an economic system which is ostensibly capitalistic, governmental action may forbid altogether the production and sale of some commodities and services, may regulate the production and sale of other economic goods, may interfere with the ordinary bargaining processes of labor markets by sponsoring and encouraging labor organizations, and may establish minimum wages and maximum hours for workers—to mention only a few things. After governmental interference and control is an established fact, the productive and distributive processes of the economy will undoubtedly adjust themselves again on the basis of price relationships in the new situation, but the total results achieved will never be quite what they would have been in the absence of governmental intervention.

Noncompetitive Conditions. Our discussion of the control of production on the basis of prices has assumed the existence of competitive relationships among buyers and sellers in the market. It is only under competition that the enterprisers in any particular industry, in reacting to increased demand and favorable price-cost relationships, may be expected to expand production so much that the price of the product will necessarily fall back to the cost-of-production level. In controlling the affairs of their enterprises, monopolists and monopolistic competitors must pay attention to the demand for their products, but under any given demand they have no incentive to carry production to the point at which they will receive prices which will only cover the average cost of production per unit of their goods. Instead, as we shall see in detail in connection with the theory of price determination, they tend under any given conditions of demand and productive capacity to limit output to the volume which will maximize their profits. Since such an output is ordinarily well short of that at which price just covers average cost of production per unit, it follows that the productive results of operation under conditions of monopoly or monopolistic competition are quite different from those envisaged by the general description of the control of production on the basis of price relationships. In somewhat similar fashion, the existence of conditions of monopoly or semimonopoly on the demand or supply side of the markets for the agents of production, or even the existence of monopoly or monopolistic competition in the sale of finished products whose producers are also, of course, demanders of the productive agents, will have an influence on the prices at which the total available stocks of productive agents can find employment, on the distribution of the agents among industries, and on the kinds and relative quantities of economic goods which the entire economic system will produce.

Under conditions of monopolistic competition, products are differentiated and the individual producer is in a position to control the market for his product to some extent. In such a situation, the enterpriser may readily decide that it will be better for him to try to "educate" the consumers to want what it seems appropriate for him to produce than to try to suit his product to the supposed wants of consumers. The simple theoretical discussion of the control of production on the basis of price relationships tends to overlook the whole range of advertising and other economic activities whose purpose is, at least in part, to change the pattern of human wants rather than to satisfy the wants which exist already.

Turning now to other functions of the price system, we note that the method of apportioning or rationing consumers' goods and services among the persons who desire them by means of price changes, or of bringing into adjustment the market demands for and market supplies of economic goods by means of price changes, is capable of working very efficiently. There is always a price high or low enough to move any quantity of an economic good, however large or small, off the market. In actual practice, however, the method often works much better in one direction than in the other. When certain economic goods are unusually scarce, prices ordinarily rise readily and the numbers of buyers or the quantities demanded are restricted. When economic goods exist in quantities which cannot be disposed of at the prevailing prices, the prices do not always fall readily, and we sometimes see large unsold stocks of these goods piling up while many people desire to purchase them and cannot do so. The difficulty here is not that the price mechanism will not work but rather that producers and sellers do not care to let it work. That is, the enterprisers, unlike the individual producers who operate under strictly competitive conditions, are able to exercise control over the prices of their products and are unwilling to lower prices to levels at which all the available goods could be sold.

The method of controlling long-run supplies of productive agents by means of prices applies with real force only to capital, and even here the results produced are far from perfect. The total quantity of savings and of capital often adjusts itself very slowly to changes in the interest rate. In times of depression saving may go on at a considerable rate even though the rate of interest has been declining steadily and the demand for funds to invest in industry is light. On the other hand, in highly prosperous times the demand for savings to invest in productive facilities may outrun the total quantity of savings even though the rate of interest rises considerably.

EQUILIBRIUM AND OPTIMUM

Equilibrium as an Economic Optimum. While full economic equilibrium is not likely to be completely worked out at any particular time, we have seen that there are forces making for equilibrium in the operation

of a capitalistic economy on the basis of price relationships. And there are some loyal supporters of the capitalistic system who contend that such an economic equilibrium would be an economic optimum as well and would result in the maximization of economic and social welfare for the citizens of the economy. As the result of the control of production on the basis of price relationships, only those goods will be produced by the members of society as producers which the members of society as consumers desire, and the relative quantities of the various economic goods produced will be adjusted to the demands of consumers for them. The production of each economic good will then be carried to the point where the last unit produced will command a price sufficient, and only just sufficient, to cover the cost and effort of producing it. At the same time, the consumption of each economic good will be carried to the point where the last unit purchased is expected to provide a degree of satisfaction which is sufficient, and only just sufficient, to justify the price paid for it. Since at this margin both expected satisfaction on the one hand and cost or effort on the other are supposed to be equal to the same thing (the price of the good), it is customary to assert that they are equal to each other. But, because other units of the good before the margin is reached are expected to give more satisfaction than the marginal unit and are expected to cost less than the marginal unit, ideally the industry which produces that particular good, and hence all industries, will control production so that they furnish society with the greatest possible surplus of satisfactions over costs.

If the agents of production are allocated among firms and industries strictly on the basis of prices, their employment will be perfectly coordinated with the desires of consumers as expressed on the market. No unit of any productive agent will be used for the production of a particular economic good if its use in the production of some other economic good would result in a greater value-product or a higher rate of remuneration for the owner of the agent. Because of the apportionment of finished goods and services among consumers on the basis of prices, no unit of any economic good would be purchased and consumed by a given person if a greater marginal satisfaction would result from its purchase and consumption by someone else (in so far as degrees of expected satisfaction are accurately measured by the willingness of individuals to pay prices on the market). Finally, in the state of full equilibrium on the basis of price relationships, the productive resources of the economy would be divided between providing for the present (through the production of finished commodities and services) and providing for the future (through the production of capital goods) in perfect accordance with the desires and preferences of the citizens. Such is the attractive picture which is sometimes offered in connection with the operation of the price system and its results.

Opposing Considerations. While any discussion of the "goodness" or "badness" of the results produced by the operation of the price system is not strictly a part of our business in investigating the principles of economics, there may be no harm in examining some of the reasons which exist for doubting that full economic equilibrium on the basis of price relationships would also be an economic optimum. In the first place, it is clear that the control of production through changing price relationships merely adjusts production, at best, to the effective demands for goods in the market, and not to basic human desires or needs. Production would be adjusted to human desires or needs only if all persons demanding economic goods had substantially equal money incomes (and perhaps not even then). Under this condition, consumers would presumably offer highest prices for the goods which they desired most. Thus the existence of a high price for a good would indicate a strong consumer desire for more of it, while low prices would prevail for unimportant goods or goods the demands for which were already being relatively well satisfied. Under the actual conditions which prevail under capitalism, with great inequality in the distribution of money income among individuals, the reliability of these price indications is partly destroyed. A high price may exist either for an essential good which is in strong demand or for some trivial, unimportant good whose purchasers have a great deal of money to spend. Conversely, a low price for a good may mean that the good is of little importance or that the demand for it is already well provided for; or it may mean that the good is one of considerable importance, perhaps even a necessity of life, and that many unsatisfied desires for it remain, but that those who desire it have little money to spend for it or anything else.

Business enterprisers, who are on the lookout for goods whose prices are favorable in relation to their costs, do not bother to ask *why* the prices are what they are. Indeed, it would be too much to expect them to investigate the reasons why people demand the goods which they do in fact demand. Thus under capitalism, with its inequality in income distribution, the reactions of enterprisers to price relationships alone sometimes serve to direct agents of production into the production of nonessential goods while really important human wants go at least partly unsatisfied. The result is that yachts are produced for the transportation of the few before there are enough shoes for the transportation of the many, and cake is produced for the rich before the needs of many poor people for bread are satisfied. Under conditions of extreme inequality in the distribution of money income, even the most perfect adjustment of production on the basis of price relationships could scarcely be considered to produce an optimum situation in terms of the satisfaction of basic human needs and desires.

In the second place, just as the money prices which individuals are willing to pay for economic goods in the market are not always reliable in-

dicators of their basic needs and desires, so the money costs which enterprisers must pay in order to produce goods are not always reliable indicators of the real costs of these goods to society. In setting up a new enterprise, the individual is likely to be concerned only with the relation of selling price to the money costs which must be paid for agents of production in the market. His enterprise may pour industrial wastes into a river with the result that the fish are killed or the water supply is polluted. Employment in his enterprise may have bad effects on the workers, and the smoke and soot which his factory emits may have unfortunate results in terms of the health of the citizens or their laundry and house-painting expenses. Such items are likely to be disregarded by an enterpriser in deciding whether it is worth while to produce a good on the basis of money prices and costs.

Thirdly, we may note that the price mechanism in the market does not furnish us with a satisfactory means of expressing all of our wants. It would be possible, we may suppose, to allow private enterprisers to furnish us with such things as fire protection, education, and health service on the basis of price relationships, but we do not do so. Because of the great importance which we attach to these services and because of our complete unwillingness to accept as appropriate the volumes of them which would be provided for us on the basis of price-cost relationships in the market, we insist that the provision of these services be taken out of the realm of the market for the most part and that it be undertaken by public authority. Again, some of our more complex wants seem to baffle the ingenuity of the market and price mechanism altogether. Take, for example, the matter of national defense, or the protection of the citizens of the economy from attack by external enemies. As one writer has said, "No device readily suggests itself by which such of the inhabitants of a given territory as were unwilling to subscribe to the cost of maintaining an army could be prevented from enjoying its protection in time of war; or, which would adjust the amount of protection enjoyed by each citizen accurately to the amount that he chose to spend upon this service." [2] Or consider the desire for economic security, which undoubtedly holds a high position in the valuation scales of many individuals. Clearly, "there is no way in which we can go into the market and, as it were, bid up the value of security, so as to stimulate the economic system to deliver more of this admirable product." [3] As a final example, suppose we have a desire to live in an economy which is marked by much less inequality in the distribution of income than is our present capitalistic system. There is apparently no way in which we can satisfy this desire through our willingness to pay prices for economic goods in the market.

[2] Barbara Wootton, *Lament for Economics*. New York: Rinehart & Company, Inc., 1938, p. 177.

[3] *Ibid.*, p. 202.

In the fourth place, the market and price mechanism, as such, gives consumers no opportunity to bid against the production and sale of certain commodities and services which they regard as undesirable. There may be a great many people whose total of satisfactions would be much increased if they could prevent the publication and sale of a particular book or the production and sale of alcoholic beverages, noxious patent medicines, or other products, and who would be glad to pay prices to obtain the satisfaction of these negative preferences if any opportunity could be given them to do so. Of course the individual can always refrain from purchasing commodities and services which he regards as undesirable, and if many individuals so refrain, the production and sale of these goods may become unprofitable and be discontinued. However, in many cases, our negative preferences can be exercised, if at all, only by public (that is, governmental) action.

Finally, the market and price mechanism itself provides no means for measuring the efficiency with which the market and price mechanism works or the total satisfaction which is derived from its use in comparison with that which might be realized from the use of some alternative mechanism for the allocation of scarce means among alternative ends. The defense of the market and price mechanism in these connections is usually stated in terms of the freedom of choice which this mechanism affords to consumers in the capitalistic system. However, on examination, we might well find that the freedom of choice which consumers would really like to have would be at the same time less detailed and more comprehensive that that which the market and price mechanism actually offers. It would be less detailed because the market and price mechanism actually offers a bewildering range of choices to consumers who in many cases lack the ability to make the necessary choices intelligently or even to determine where they go wrong so that they can correct their decisions in connection with later purchases. It would be more comprehensive because, in actual practice, the consumers' choice as to the various uses to which the productive resources of the economy should be put is secondary rather than primary. That is, the market and price mechanism never asks the consumers to specify the various commodities and services for the production of which they would like to see the scarce resources of society used. The really fundamental choices are made by business enterprisers, who decide what commodities and services should be placed on the market, and consumers can actually choose only among the various options which are offered to them by the enterprisers.

After considering all of these factors, we may well decide, as many other economists have done, that, even if full equilibrium on the basis of price relationships could be achieved under capitalism, it would be just equilibrium and nothing more. Or, if we are determined that such an equilibrium must be an economic optimum of some sort, we should con-

sider it only as a practical optimum, rather than an "optimum optimum," in the sense that market equilibrium on the basis of price relationships, with all its imperfections, would furnish us with a better set of solutions for the important economic problems than would any alternative which is also deemed practicable. However, such an opinion should be formed, if at all, only after considering any other mechanisms which exist for the making of basic economic decisions.

THE PRICE SYSTEM IN NONCAPITALISTIC ECONOMIES

The discussion of the functions of prices and the significance of the price system as a whole has thus far dealt only with the conditions which prevail under capitalism and would not be valid for socialism, communism, or fascism. Indeed one useful way of distinguishing other economic systems from capitalism is to consider the functions which prices or price relationships would be permitted to perform in the various systems. It will be remembered that the term "socialism" refers to an economic system in which the ownership and operation of land and capital are vested in society as a whole rather than in private individuals. Such a system would probably make use of money, credit, and prices, but the operation of the system would not depend upon price relationships to as great an extent as does that of a capitalistic system.

The Control of Production under Socialism. Under socialism, as the system is set up in theory, decisions as to the kinds and quantities of goods which should be produced would not depend upon price relationships. Thus under socialism one of the major functions of prices under capitalism would be eliminated. Society as a whole would be the only business enterpriser of any great importance. Since society as a whole would both own and operate the land and capital of the system, interest and rent would not be paid and received as under capitalism. The planning authority or commission, which would be entrusted with the direction of economic activity under socialism, might charge itself with certain amounts of rent and interest for the land and capital used in various branches of production, but any such charges would be purely arbitrary and would not result from competition between private owners of these agents and persons desiring to use them in production. Wages of various kinds would constitute about the only money expense of production in the ordinary sense under socialism. Wage differentials themselves would probably be adjusted so as to accomplish whatever results the planning commission thought were worth achieving, and would not result from competition between workers and private employers. Thus the cost per unit of any economic good would be whatever the planning commission said it was.

Similarly the prices of finished products would not be determined by competition between buyers on the one hand and private sellers, organized

under competitive conditions or those of monopoly or monopolistic competition, on the other. The prices of the various economic goods might well be set low or high by the planning commission according to whether they wanted to encourage or discourage the consumption of the goods or whether the amounts available of the goods were large or small. Since both prices and costs would be rather arbitrarily determined by the same planning authority, the relationships which existed between such prices and costs would be virtually without significance. That is, the economic planners, having established prices and costs at their pleasure, could not then logically depend upon the relationships between these prices and costs for guidance in determining the kinds and quantities of economic goods which should be produced. Thus while prices and costs would be used as accounting devices under socialism, and as convenient tools for making plans and supervising their execution, decisions as to the kinds and quantities of economic goods to be produced would be made by the planning commission on the basis of the presumed social need for the products, or on some other basis, and not on the basis of the prospective profitability of producing the various kinds of goods.

The Allocation of Productive Agents under Socialism. Since society as a whole would own and operate the land and capital under socialism, these agents of production would be distributed among the various industries and fields of production in such a way as to make possible the carrying out of the plans set up by the planning authority. Since the various industries would never be called on to bid various prices for the land and capital, the distribution of these agents among industries would be in no way dependent upon the prices which these industries might be willing to pay for the agents. In other words, land and capital would be assigned arbitrarily to the various industries on the basis of the plans for production, and the second function of prices under capitalism—that of allocating the available quantities of agents of production among industries—would be eliminated under socialism in so far as land and capital were concerned.

However, even under socialism the workers of the system would probably not be treated so cavalierly. Workers are human beings and the individual workers have different abilities, find some occupations pleasant and others distasteful, prefer to live in different places, and so on. Their arbitrary assignment to certain tasks might well make for both inefficiency and dissatisfaction on the part of the workers. In all probability, then, the planning authority would provide for the payment of different rates of wages in different occupations and allow the available workers to distribute themselves among the various occupations on the basis of these differential wages. Low wages would be used to prevent too great a concentration of workers in the easy, safe, and pleasant occupations, while high wages would be used to insure the presence of at least a minimum number of

workers in difficult, dangerous, and distasteful jobs. The wage differentials used might not be nearly so large as under capitalism, and, instead of being determined for the most part by the competition of workers and employers as under capitalism, they would be set up and varied at will by the planning authority in order to obtain the desired distribution of workers among industries and occupations. However, in spite of these considerations, it may be said that under socialism the distribution of workers among industries and occupations would depend to some extent upon price (wage) relationships.

The Rationing of Consumers' Goods under Socialism. A socialistic system would also face the problem of rationing the limited quantities of finished goods among the consumers of the system. It is likely that this function would be performed for the most part by means of price changes under socialism as under capitalism, although the necessary rises and falls in prices would result from orders given by the planning authority and not from the actions of individual buyers and sellers in the market. If, at a certain time, the available quantity of an economic good should be unusually limited, the planning authority would raise the price of the good so as to eliminate some buyers from the market. On the other hand, a low price would be used to induce increased purchases of some economic good which was available in unusually large amounts. As a matter of fact, this method of rationing might work better under socialism than under capitalism, since the planning authority probably would not hesitate to raise or lower individual prices by any amount necessary to bring market demand and market supply into equilibrium. However, under socialism, some rationing might well be done through quantities instead of through prices in order to make sure that all consumers would be able to obtain at least certain minimum amounts of scarce commodities.

The Control of Capital Formation under Socialism. The total amount of land available for production in a given economic system would be as fixed under socialism as under capitalism, and the total population and labor supply would probably be as little subject to wage influences under one system as under the other. However, a socialistic system would have to find some method of reaching a decision as to the extent to which existing agents of production should be used to produce capital goods to assist in further production instead of being used to produce consumers' goods and services for present enjoyment. Under capitalism, the processes of saving, investment, and capital formation are allowed to depend to a considerable extent upon the reactions of individuals and firms to changes in the rate of interest, but under socialism this function of price changes would probably be eliminated.

That is, the planning authority would decide to what extent capital goods should be produced, in order to achieve the greatest good for society

as a whole or some other objective, and it would accomplish its will in this matter simply by assigning existing land and capital to the production of capital goods rather than consumers' goods, and by paying high enough wages to attract workers to the industries producing capital goods. Under capitalism, the individuals who save are in general those who wish to do so on the basis of the interest rate and the strength of their desire to consume their incomes in the future instead of in the present, and only those persons who save receive interest when the resulting capital goods become effective in production. Under socialism, all persons would save by going without consumers' goods, for each increase in the quantity of productive agents used to produce capital goods in a given short period involves a corresponding decrease in the quantity of productive agents which are available to produce consumers' goods in that period. Similarly, when the resulting capital goods became effective in production, all persons would share in the benefits of increased production by receiving higher real wages.

Thus society as a whole saves under socialism and enjoys the benefits which result from the use of increased amounts of capital in production, while individuals save under capitalism and benefit by so doing. Under capitalism individuals decide how much or how little to save, while under socialism this decision is made by society as a whole through the planning authority. However, the process of saving and capital formation is basically the same in the two systems, because the formation of capital goods depends upon saving in any system, and saving always involves going without certain amounts of consumers' goods, which might otherwise have been enjoyed at present, in order to secure larger amounts of consumers' goods in later periods of time. The difference between the systems is found in the dependence or lack of dependence of saving and capital formation on price relationships.

The economic system of Soviet Russia is far from being a perfect example of a modern socialistic system, but it is more nearly socialistic than anything else. Consequently its basic economic decisions are made at least roughly in the fashion which has been outlined as probable for the ideal theoretical socialistic system, except that the apportionment of finished goods among consumers has been accomplished to a considerable extent thus far by allotting physical quantities of goods to various classes of consumers rather than by raising and lowering individual prices.

Making Economic Decisions under Communism. The functions of prices or price relationships under capitalism would be partly retained and partly eliminated under socialism, as we have seen, but they would apparently be entirely eliminated under ideal communism, for such a system would try to get along without using money and prices at all. Decisions as to the kinds and quantities of goods to be produced, as to the appropriate allocation of land and capital among industries and fields of production,

and as to the extent to which saving and capital formation should go on, would be made, under communism as under socialism, without benefit of price guidance. In addition, under communism, workers would be expected to distribute themselves among industries and occupations in accordance with society's need for their services and without any differences in wages existing between occupations. Moreover, communism looks forward to the production of various commodities and services in such abundance that no system of rationing, by price changes or otherwise, would be necessary. Each person would receive real income according to his needs and would be able to acquire the things which he needed by helping himself to them at various public storehouses or centers of distribution.

Making Economic Decisions under Fascism. As we have seen, the fascist economies of Italy and Germany had all the institutions of capitalism at least in name. The leaders of the economies said at the beginning that their governments did not intend to direct and control economic activity in the long run but had in mind merely "the opening up of the way" for private industry and trade. The leaders loudly proclaimed their faith in private initiative as an effective and economical instrument for getting things done, although they held that, since the interests of the state were of supreme importance, the individuals who directed economic activity were responsible to the state. The state was to interfere in economic activities only when private initiative was lacking or inadequate, or when the interests of the state were so directly involved as to make governmental control necessary. Governmental interference, when it occurred, might take the form of encouragement, general supervision, or outright management.

Under these conditions, all the important economic decisions made on the basis of price relationships under capitalism could be made in the same general way under fascism except when the state stepped in and controlled these decisions in its own interest. Individuals could decide what to produce and in what quantities unless the state decreed that certain lines of production should be speeded up, slowed down, or discontinued altogether, or that the production of new kinds of economic goods should be undertaken. The distribution of land, labor, and capital among industries occurred on a price basis unless the state saw fit to interfere with the process. Individuals were free to save or not to save unless the state compelled them to buy government bonds, took their incomes by taxation in order to be able to make certain grants of capital funds to industries which it wished to encourage, or required corporations to reinvest earnings instead of paying them out in dividends to stockholders. Finally, in performing their customary rationing function, the prices of economic goods could rise or fall unless the state stepped in to fix or control prices, as it often did.

However, as the fascist economies actively prepared for war and finally became engaged in war itself, their governments found it necessary to in-

terfere with economic activity to an ever-increasing extent. In the end, practically all important economic decisions were made on the basis of governmental authority rather than on that of price relationships, and the fascist economies could be said to operate on the basis of national economic planning of a sort. The fascist governments achieved their control over economic activity by superimposing various agencies and regulations on what would otherwise have been capitalistic economies, and not by owning and operating the industries themselves, but the significance of price relationships declined rapidly as governmental control over economic activity increased.

QUESTIONS AND PROBLEMS

1. "Rises and falls in the prices of finished economic goods and of the agents of production necessary to produce them control the kinds and quantities of economic goods which will be produced under capitalism." Explain.

2. Suppose a piece of land becomes vacant in the downtown section of a city. What, under capitalism, will determine whether this piece of land will be used as a site for a factory, office building, hotel, theater, garage, department store, or for some other purpose? How would your answer be different if the prevailing conditions were those of socialism?

3. How are capital funds distributed among the various industries and firms which are anxious to obtain them under capitalism?

4. A few decades ago there was no radio industry in the United States. Now it employs many thousands of workers. How does such a developing industry obtain its labor force under capitalism?

5. Limited supplies of finished economic goods may be rationed among consumers either through price changes or by limiting the quantities which individual buyers may purchase in a given period of time. Compare these two methods of rationing as to efficiency and effects on consumers.

6. "Price changes are relatively impotent in controlling the total quantities of land and labor in the long-run." Why?

7. "The total supply of capital funds depends entirely on price relationships in the long run." Show whether you agree.

8. What are some of the obstacles to the attainment of full economic equilibrium on the basis of price relationships under capitalism? Explain.

9. "If our capitalistic system operated perfectly on the basis of price relationships, the result would be not only economic equilibrium but economic optimum." Show whether you agree.

10. What is the significance of great inequality in the distribution of income in connection with the question of whether economic equilibrium on the basis of price relationships would also be an economic optimum? Explain.

11. "The price mechanism under capitalism is defective in that it does not permit the expression of certain negative preferences of individuals nor of certain complex positive wants." Explain.

12. "Under socialism, some functions which prices perform under capitalism would be retained, while others would be modified or eliminated." Explain.

13. "Decisions as to the kinds and quantities of goods to be produced under

socialism would probably be made arbitrarily rather than on the basis of price relationships." Show whether you agree.

14. "Under socialism, land and capital would be distributed among industries in one way while workers would be distributed in a quite different way." Why?

15. Compare the process of saving and capital formation under socialism with that which exists under capitalism.

16. "The functions of prices would be completely eliminated under communism." Explain.

17. "The price system had the same significance under fascism as under capitalism." Show whether you agree.

REFERENCES FOR FURTHER READING

Blodgett, R. H., *Comparative Economic Systems.* New York: The Macmillan Company, 1944, Chaps. 5–6.

Dixon, R. A., *Economic Institutions and Cultural Change.* New York: McGraw-Hill Book Company, Inc., 1941, Chaps. 5–8.

Wootton, Barbara, *Plan or No Plan.* New York: Rinehart & Company, Inc., 1935, Chaps. 1 and 3.

V

Characteristics of Production

PRODUCTION IN GENERAL

The Nature of Production. We have already noted in Chapter II that the people who engage in the transporting, storing, and marketing of economic goods, as well as those who render direct personal services to other people, should be classed as producers along with those who actually "make things." This conclusion as to the nature of production and producers would not have been accepted by the economists of many years ago, for their concept of production long included only the making of material or tangible objects. Workers in the fields of transportation, storage, and marketing were regarded as estimable and useful citizens, but they could not, it was thought, be classed as producers because in the physical sense they produced nothing. In similar fashion, doctors, lawyers, teachers, actors, churchmen, and other people who satisfied human wants through direct personal services were classed as nonproducers because, as Adam Smith once suggested, their labor does not "fix or realize itself in any permanent subject, or vendible commodity."

Later on it came to be realized that no producer really creates any physical objects or adds anything to the sum total of matter already in existence. Even the workers who "produce" material articles merely take something which is already available and change its form in such a way that its presumed usefulness to someone, and hence its wantedness or desiredness, is increased. As soon as this fact was realized, it was a relatively short step to the conclusion that all workers who create utility or add to the wantedness or desiredness of goods are in reality producers. Thus production is usually defined in economics today simply as *the creation of utility,* and the term includes the activities of many people who do not turn out tangible articles.

Types of Utility Creation. The most obvious type of production or utility creation is found in what is called the creation of *form utility*. The creators of form utility include all persons who engage in adding to the utility

of tangible articles by changing their form. People who cut down trees, make them into lumber, or use the lumber to turn out chairs, tables, beds, and other desired articles, qualify as creators of form utility, as do other people who raise steers, make hides into leather, or convert leather into shoes, gloves, pocketbooks, brief cases, and other things. These and similar activities are so readily accepted as productive by most people that the multiplication of illustrations seems futile, but it should be emphasized that the creation of form utility occurs at all stages of the production of commodities. People who produce raw materials, or machines for raising or extracting raw materials, create form utility just as much as those who put the finishing touches on articles which are almost ready for the final consumers.

However, in addition to being physically ready for use, articles must be available at the time at which they are desired by the users. Thus people who hold and store commodities and dispense them as they are required by purchasers are considered to be creators of *time utility*. People who operate cold-storage plants or grain elevators clearly perform this function. Eggs, for example, are put into cold storage in seasons of the year when they are plentiful and relatively low in price, and they are taken out and sold at other times when fresh eggs are scarcer and relatively high in price. The fact that people will pay more for cold-storage eggs later on than they would have paid for these same eggs when fresh indicates that utility or wantedness has been created by the storage process. Wholesalers, jobbers, retailers, and other middlemen also create time utility to some extent. Storekeepers, for example, purchase articles in relatively large amounts, hold them until they are required by consumers, and sell them to consumers in small, convenient amounts. Speculators in produce are also sometimes classified as creators of time utility. By buying and holding, say, wheat when they think its price will go up, or by selling short when they foresee a decline in price, they bring about a change in the time at which certain amounts of the available stock of wheat will be consumed. If their decisions are correct, they may be said to create time utility.

It is as important to have articles available at the desired place for consumption as it is to have them ready at the desired time or to have them made in the first place. Hence all persons who work in transporting articles to bring them to or nearer to the place of consumption may be said to create *place utility*. This type of utility creation covers a wide range of activities. It includes the delivery of groceries by the retailer's truck to the consumer's house as well as the moving of oranges by railroad from California or Florida to Illinois, or the transportation of jewelry by airplane from New York to St. Louis. Even arbitragers, who make an income by purchasing things in one place and selling them immediately in another place, are sometimes classed as creators of place utility.

A fourth type of utility creation in connection with tangible com-

modities may be distinguished. Commodities which are made, stored, and transported must be transferred into the hands of the final consumers before production is completed and consumption may begin. A radio or a suit of clothes is incompletely produced as long as it is in the hands of a storekeeper or dealer and it has greater utility or wantedness for the consumer than for the storekeeper. Consequently, we need to credit the storekeeper or other seller with the creation of *possession utility* when he sells the article in question to the consumer. Some individuals, such as salesmen, may devote practically all of their efforts to facilitating the transfer of commodities from producers to consumers, but in other cases possession utility is created by individuals, such as retailers or manufacturers, who also create time, place, or form utility.

The types of utility creation thus far discussed are all carried on in connection with tangible commodities and consist in changing the form, place, possession, or time of consumption of these commodities in such a way as to increase their utility. The final type of utility creation consists in the rendering of personal services which are directly desired in and of themselves for the satisfaction of human wants. Thus the doctor who cares for me during an illness and the nurse who ministers to my comfort, the lawyer who defends me against the legal claims of other individuals, the barber who cuts my hair, the professional baseball players, actors, and singers who entertain me, the educator who instructs me, and many other individuals, are known as creators of *service utility*. However, it is worthy of note that some of the services of such people as doctors and lawyers are performed for business firms in connection with the various processes of commodity production and not directly for individual consumers as such. For this reason it may be said that some, at least, of the creators of service utility also assist in the various other types of utility creation.[1]

"Unproductive" Activities. Some economists, in writing about the nature of production, have felt it necessary to describe some economic activities as "unproductive" in character. In reaching the conclusion that some economic activities are unproductive, they usually shift, consciously or unconsciously, from the definition of production as the creation of utility to an ethical concept of production as something which enhances the "welfare of society." The arguments which rage over modern advertising are a good illustration. Much of modern advertising, we are told, is not educational or informative in character. It is not designed primarily to inform people

[1] Workers such as stenographers, office boys, janitors, filing clerks and night watchmen, who do not handle or deal with the articles which their firms produce, are nevertheless usually classified according to the type of utility creation carried on by the firms which employ them. Thus a stenographer would be said to create form utility if she worked in a factory office, place utility if she worked in the office of a railroad company or trucking concern, and so on. On the other hand, bankers, insurance workers, certified public accountants, advertising workers, and others are often classified as general auxiliary types of utility creators, since they are likely to give assistance to all the various standard types of utility creation at one time or another.

about new products or new uses for old ones, or to tell them when and where certain goods will be available. Instead, a large part—some say nine tenths—of advertising today is competitive or combative. It seeks to divert people from one brand of good to another or from one seller to another.

Thus cigarette advertising today urges us to forsake all other brands and concentrate on smoking Lucky Strikes, Camels, Chesterfields, Old Golds, or Philip Morris. Of course such advertising may actually induce some people who would not have smoked otherwise to take up and enjoy the smoking of cigarettes. In other cases people may shift from one brand to another from which they derive a greater amount of genuine enjoyment, but in general the advertising is deemed unproductive and the advertisers are compared to a group of gamblers sitting in a poker game with a fixed total amount of funds among them. Some players may gain and others may lose, but the group as a whole cannot gain. Similarly, if there is no gain in the total number of smokers or in the total satisfaction derived from smoking, the cigarette companies are operating in a closed market. Some may gain customers by advertising, while others lose customers, but the group as a whole is no better off than it would have been if the advertising activities had not taken place.

From the point of view of social welfare it is argued that the resources of production which are devoted to such competitive or combative advertising are wasted. They not only accomplish nothing which is useful to society, but our society as a whole would gain if the workers and other agents of production now used in this type of advertising could be shifted to some other employment in which they could confer a benefit on society. Now there may be some merit in this argument from the point of view of social welfare. In view of the host of unsatisfied human wants which still exist, it certainly cannot be answered by asserting that the workers and other agents of production now used in advertising would have to stand idle if advertising activities were discontinued. However, the argument does not furnish a satisfactory basis for eliminating advertising activities from the category of production. Production is not defined as activities which enhance the social welfare but rather as the creation of utility or wantedness. The facts in this case seem to be that many people want the advertising activities to be carried on and are willing to pay to get them carried on; hence we must admit that these activities belong in the field of production. And this is true in spite of the fact that advertising services are demanded by businessmen and firms and not directly by the final consumers. The services of workers who produce factories and machines are also demanded by businessmen and firms and not directly by the final consumers, but we do not consider the workers' activities unproductive for that reason.

Other writers would class as unproductive all labor and other re-

sources devoted to operating houses of prostitution, gambling dens, bookmaking establishments for betting on races, and all facilities which people may use for gambling on changes in the prices of securities or commodities. They argue that the welfare of society as a whole would be enhanced if people did not desire that such activities be carried on, or if they were compelled to get along without having these desires satisfied. We do not want to take the attitude that all or any of these activities are morally or socially desirable, but we must contend that such questions of ethics or morals are outside the strict province of the economist as a scientist. On the simple question of productive and unproductive activities we must decide, according to the definition of production, that these activities are productive since *some people* want these things done and are willing to pay to get them done. Again, in the case of soldiers, policemen, judges, lawyers, and other workers, we may concede that our society might well be better off if we did not need the activities of such people, but apparently we do need them, or at least we want them and are willing to pay for them.

But, we may be asked next, what about such activities as the production and sale of those patent medicines which are useless, if not downright harmful, when employed in the treatment of the ailments or diseases which they are supposed to cure? What about people who produce and sell to an unsuspecting public adulterated foods or shoddy clothes at prices high enough to cover the costs of producing and selling first-class commodities? Without condoning these activities in any way we must contend that we cannot eliminate activities from the field of production merely because the resulting products do not give as much satisfaction to the purchasers as these people expect to get when they buy the articles. Even when fraud and misrepresentation are not involved, such experiences are common among the consumers of commodities. There would be little left of the concept of production if we considered as unproductive all activities whose resulting products did not give the consumers as much satisfaction as the consumers expected to get. As private citizens we may condemn fraud and misrepresentation in unmistakable terms, but as economists we cannot allow these practices to cancel production so long as people want the goods in question and are willing to pay for them.

Finally, we may be asked, if the usual definition of production compels us to consider many socially undesirable activities as productive, why do we not abandon this definition and turn to one under which production would include only those activities which enhance the social welfare? The most compelling reason for not following this suggestion is that the economist has no scientific method for determining, on behalf of all the people, just what social welfare is. We should be compelled to decide, in following this suggestion, that some of the ends which people seek are bad while others are good, and that some of the wants which people have should

be satisfied while others should remain unsatisfied. Such decisions are outside the province in which the economist as a scientist must work. He must take the ends which people seek as given, and study how scarce means with alternative uses are allocated among these ends. As individual social philosophers, we are at liberty to decide that some productive activities are socially undesirable and even to work to get them discontinued, but as economists we must conclude that all activities which people desire to have carried on and for which they either actually pay or would be willing to pay are productive.

Even if we did want to step out of character and define production in terms of the enhancement of social welfare, we should find this concept of production to be of comparatively little use. Different people have quite different opinions on ethical and moral questions, and very few people would probably agree with us exactly as to the things which should be included in and excluded from production. In effect, each individual would have to decide for himself just what activities should be classed as productive on the basis of their contribution to the social welfare, and there would be almost no limits on the types of activities which some people would have us exclude from the category of production. Certainly a considerable number of people would want to class as unproductive all the labor and other agents of production now employed in turning out alcoholic beverages. Others would want to extend the unproductive category to include the industries which turn out tobacco goods and playing cards, and strict vegetarians might conclude that the social welfare would be enhanced if no one raised animals for meat or engaged in meat packing. Our conclusion, then, is that the concept of production, if it is to be useful in economic analysis, must not rest on ethical considerations.

DIRECT AND INDIRECT PRODUCTION

The Agents of Production. Production, or the creation of utility, always requires some combination of the agents of production. Since the agents of production have been described in another connection, no more than a few additional considerations will be taken up here. Only land and labor are ordinarily considered as primary agents of production. Land, it will be remembered, includes all natural wealth, or wealth which is not produced by human labor. It is indispensable to all branches of production, but labor must be applied to it before commodities and services become available for consumption. Labor is defined as *human energy which is expended for the purpose of acquiring income.* The income which is referred to here may be money or real income, but not psychic income. Human exertion which is undertaken merely for the pleasure or satisfaction involved directly in the activity is not usually classed as labor. If a man goes mountain climbing for exercise or to obtain a beautiful view of the surrounding coun-

try, he would not be classed as a laborer because of this activity. However, he might retrace his footsteps the next day as the hired guide of another party of mountain climbers and then be considered a laborer even though he expended no more human energy than on the other occasion. It is also true that the connection between the expenditure of human energy and the resulting income must be fairly direct and immediate. A college football player would not be considered a laborer merely because he hoped that after his graduation some rich alumnus, remembering his feats on the gridiron, would offer him a position as a bond salesman or an insurance agent. On the other hand, a college football player who played for money wages or direct real income would be a horse of another color.

Much discussion has centered around the question of whether business enterprisers should be classed as laborers or should be distinguished as a separate agent of production. Those who hold the latter view argue that, while business enterprisers are laborers in a broad sense since they expend human energy for the purpose of acquiring income, their functions differ so widely from those of ordinary laborers that they deserve to be set apart. Ordinary workers perform a task for wages, but business enterprisers must assume ultimate responsibility for the operation of a business. They must see the opportunity for the operation of a new enterprise or concern, contrive to obtain funds to establish it, attract other necessary agents of production, organize and combine them into a going concern, undertake to pay these agents at specified rates, and depend upon the successful outcome of the business for their own income, if any. When we are told that the business enterpriser needs to be bold, fearless, courageous, quick in making decisions, shrewd in bargaining, well informed as to business conditions, and full of executive capacity, powers of organization, and keenness in judging men, we realize that the enterpriser must be a giant among men.

However, although there are undeniable differences between the functions involved in organizing and operating a business and those of ordinary workers, it seems better for many purposes to class business enterprisers as laborers and to use the term "labor" to mean all human energy expended for the purpose of acquiring income, whether in mental or physical work. As a matter of fact, in the case of the modern corporation, many of the functions which were once performed by individuals who both owned and operated business units are now delegated to hired workers or managers, while the actual owners who assume ultimate responsibility for the operation of the corporate business (the stockholders) have very little to do with its actual operation or management. Moreover, even those economists who consider business enterprisers as a separate agent of production later in their books often classify these same individuals in one or more of the so-called

noncompeting groups of *labor* and discuss the determination of their ordinary remuneration in the chapter on *wages*.

Direct and Indirect Production. Capital consists of produced objects of wealth which assist directly in further production and yield satisfactions to their users only incidentally if at all. Since it results from applications of labor to natural wealth (with the assistance of the capital already in existence once the process is under way), by its very nature capital is a secondary or derived agent of production and not a primary agent. The distinction between direct and indirect production turns on the question of whether or not capital is used in the processes of production. *When consumable goods are secured by the application of labor directly to land or natural wealth, production is said to be direct.* Since almost no direct production occurs today, it is necessary to consider primitive economic conditions in order to see what direct production would be like. We might, for example, think of a group of savages living in the wilderness. They could perhaps obtain some food by picking wild fruits and berries, drinking water by scooping up water from the river with their cupped hands or by lying down on the bank to drink, and shelter by using human labor alone to make natural caves more habitable. In other cases, direct production would be more difficult. If the savages desired to vary their diet with fish, direct production would require them to lie in wait beside the stream until some unsuspecting trout swam idly past and then catch the fish by plunging hands and arms rapidly and accurately into the water. Needless to say, the standard of living of the people would be low on the basis of direct production.

Production in which capital goods are used is called indirect. In this type of production, the primary productive agents are used first to produce capital goods and only later, assisted by these pieces of capital, do they go to work to turn out consumable goods. Indirect production requires that a considerable amount of preparatory work be done before goods directly capable of satisfying human wants are produced. In other words, it means that the primary agents of production—land and labor—are used to bring about relatively remote results in terms of consumable goods, instead of relatively immediate results. Since the period of time which must elapse in indirect production between the original application of labor to land and the enjoyment of consumable goods is often quite long, this type of production is frequently called roundabout.

In observing how direct production may give way to indirect, we may return once more to our savage fishermen. It is reasonable to suppose that the savages would show considerable variations in individual efficiency in the direct production of fish. Some might become so skillful in catching fish that they could produce, with one full day's labor, enough fish to provide for two or three days' consumption. The time advantage thus acquired might be wasted in idleness or used for the direct production of

other consumable goods, but some of the more enterprising savages might use it to construct crude fish nets out of tree branches, the tendrils of vines, animal tendons, and what not. With such capital or producers' goods the efficiency of production would probably be increased considerably, but the repeated construction of nets would eventually leave these original capitalists with more capital goods than they could themselves use. Then some of the less skillful fishermen might ask for a loan of some of the fish nets so that they too could catch more fish; whereupon the owners of the nets might express a willingness to lend their nets only on condition that the borrowers turn over to them a portion of each day's catch for the use of the nets. Such payments to the capitalists could readily be construed as interest.

Of course, the more efficient fishermen, instead of constructing the nets themselves, might have agreed to fish day after day and support other members of the tribe while the latter workers were making the nets. The fact that the nets would then belong not to these workers but to the more efficient fishermen would give any socialists in the group a chance to say that the owners of the nets had exploited the workers who made the nets, since the owners would from that time on have their incomes increased by using capital goods produced by the labor of other people. The net owners would probably point out in reply that they had been compelled to give up their leisure time and surplus fish in order to support the producers of the nets during the construction of these capital goods, since the net producers could not eat their own product, and that they had been compelled to wait until the nets had been constructed and put into operation before enjoying an addition to their incomes. They would also assert that the net makers could have owned the nets themselves if they had been able to catch enough fish to support themselves while they were making the capital goods. However, the socialistic savages would contend that the best method of securing the capital goods would have had the tribe as a whole support the workers while they were making the nets. The capital goods would then have belonged to the whole community and all the citizens would have shared in the benefits resulting from their use later on.

Saving and Capital Formation. Before leaving the savages to their controversies, we may say that, although this illustration is purely imaginary, it does serve to direct attention to some important considerations in connection with saving, capital formation, and the roundabout process of production. Besides showing how capital and interest might come into being, the illustration suggests that saving and capital formation can take place only when surplus income is available above the current needs for subsistence. From the physical point of view, if the people of an economy

consumable goods to live on, the prospects of saving and capital formation are poor indeed. From the financial point of view, the same thing is true if the people must spend their entire money incomes for consumable goods. However, if they can produce surplus goods over and above present needs for consumption and if they have money incomes which more than suffice for their immediate consumption needs, saving and capital construction can take place, productivity increases, a greater surplus is produced, more capital goods can be turned out, and so on.

Surplus income and saving are necessary to capital formation, but savings must also be invested before the roundabout process of production can be furthered. If the more efficient savage fishermen had used their surplus fish merely to provide themselves with more extensive consumption than usual or with two or three days of loafing, the fish nets would not have been made. In a money economy, if people create savings only to hoard them in the form of cash or to spend them later on for consumable goods, capital formation does not take place. In similar fashion, if our savings are used to buy land, houses and buildings already constructed, or corporate securities (stocks and bonds) in which other people have previously invested, no new capital goods are created unless the people who receive our savings use them to make new investments. Those who save may invest their funds directly, or indirectly through a variety of agencies, but invested they must be if capital formation is to take place.

In our primitive illustration the individuals with the surplus income and savings set about directly to produce capital goods, but this does not usually happen in our modern economy. Instead, our savings take the form of money and are used to pay the workers and other agents of production which are used in the production of capital goods. Since the roundabout process of production is time consuming, workers must be paid, materials must be purchased, and many other costs must be met long before the time when the final consumable goods become available. The invested savings of persons with surplus incomes are used to make advances to the various productive agents which are used in the preparatory stages of the roundabout process, and the savers must wait until some future time when final products become available in order to get a return on their investments.

Our illustration also suggests a principle, applicable in any type of economic system, with regard to the cost of capital formation. In any society in which the means of production are scarce in relation to the alternative uses which we wish to make of them, the cost of obtaining capital goods, which will give us a more abundant life in the future, is found in the necessity of getting along for the present without certain quantities of consumable goods which we could otherwise have had. That is, capital goods can be obtained only by devoting to their production

certain quantities of the productive agents which could otherwise have been used to secure consumable goods for us at the present time. The quantity of capital goods which it is thought worth while to secure, the ownership of these goods, and the identity of the people who are to enjoy the enlarged future income may vary as between capitalism, socialism, communism, or fascism, but the nature of the cost of obtaining capital goods remains the same in all systems.

Modern Roundabout Production. The production of fish is made slightly roundabout when crude fish nets are constructed, but the modern fisherman approaching the stream armed with steel rod, reel, hand net, fish basket, rubber boots, hooks, lines, and sinkers represents the culmination of a process which is tremendously more roundabout. Before he can function, iron ore must be taken from the ground, made into pig iron and later steel, and used not only as a material in producing rods, reels, and hooks but in producing machinery to be used in fabricating these products and machinery to make the machinery which produces the final products. In fact, the production of iron and steel requires the use of many types of equipment also made of iron and steel produced in still earlier stages. The lead used in making sinkers also is the result of a long and complicated extractive and manufacturing process. The hand net requires not only a metal frame but many strong strands of textile material. It can be produced only if cotton is raised, harvested, separated from the seed, spun, and made into strong cord.

Rubber for the boots may be imported from the Far East or produced synthetically in the United States by a complicated process which depends on the previous production of petroleum or alcohol. Silk for the fish lines may be imported from the Far East and cork for the handle on the rod from sunny Spain or Portugal. The imports are brought to this country in ships which also had to be produced, along with the hundreds of types of equipment which they require, and goods have to be produced for export if imports are to be obtained in the long run. The various pieces of fishing equipment are produced in factories, and the construction of factories requires that rocks be quarried, bricks produced, and cement manufactured. And, of course, in all of these processes of production, tremendous quantities of human labor are required.

This illustration could be spun out indefinitely, but enough has been said to indicate the extreme complexity and roundaboutness of the productive process which lies behind the modern fisherman. And the process by means of which we obtain almost every commodity which we consume today is similarly roundabout and complex. Primitive man introduced some roundaboutness into the production of water when he managed to fell trees and scoop out rude buckets for carrying water, instead of repairing to the stream whenever he wanted a drink; and into the production of food when

he made crude baskets in which to collect and carry natural produce. However, a full description of the processes which we use today to obtain the water we drink and the food we eat would require pages of print. Increasing roundaboutness or indirectness of production has been an outstanding characteristic of our advancing economy. From the point of view of individual industries, increasing roundaboutness has meant a great increase in the amount of preparatory work which must be done before final consumable goods can be produced, and hence a similar increase in the period of time which must elapse between the original utilization of the primary agents of production and the obtaining of consumable final products. From the point of view of the whole economy, increasing roundaboutness has also meant that, with the passage of time, more and more industries have been going over to the highly roundabout process of production and that greater and greater total amounts of labor and land have come to be used to obtain relatively remote results in terms of consumable goods.

"Benefits" of Roundabout Production. Since the roundabout process of production is time consuming and produces relatively remote rather than immediate results, we must see why it is so generally used today. In the first place, after it finally gets under way, roundabout production enables us to produce greater quantities of economic goods with a given expenditure of the agents of production (or the same quantities of economic goods with smaller expenditures of these agents). When translated into money terms, this means, of course, that the economic goods cost less per unit. Again, thoroughly roundabout methods of production often turn out goods which are of better quality than those which can be secured by direct or only slightly roundabout methods. However, this is not always true. Many people consider handmade suits of clothes, for example, to be superior to those which are produced by more highly roundabout methods, and the difficulties which would be involved in producing such things as oil portraits by machine methods are obvious. Finally, roundabout production furnishes us with a number of goods which could never be produced at all by direct or almost direct methods. Automobiles, typewriters, cash registers, ocean liners, and a host of other things would probably be denied us if highly roundabout methods of production could not be used.

These three points which we have just been considering are often called the "benefits" or "advantages" of roundabout production. However, these terms imply that roundabout production is good, and we must examine the basis on which this conclusion may be reached. Actually it involves nothing more than a decision that it is desirable to maximize production and to make the scarce means of production go as far as possible in attaining the ends which we see fit to seek. If we are willing to make this decision, even though we realize that the maximization of production

may not always be synonymous with the maximization of human welfare, then roundabout production is seen as desirable and as having certain advantages or producing certain benefits.

The Extension of Roundabout Production. Although the gains from indirect or roundabout production are great, it is not possible in a given state of technology to increase indefinitely the quantities of productive agents which are devoted to securing relatively remote rather than immediate results unless we are willing to accept diminished added gains from the use of the roundabout process. Suppose two towns are located a few miles apart and their inhabitants desire to trade and communicate with each other. Under the direct method the citizens of either town would set out across country on foot to reach the other town, but communication would be slow and transportation arduous and inefficient. As a result the citizens might soon build a road between the two towns and then proceed to trade and communicate with each other. This method, while roundabout and very slow in getting under way, would eventually produce more satisfactory results than the direct method.

Once the road is built, how could more labor and land be used to further the stated objectives? A second road could be built between the two towns but, although it would be far from useless, the second road could hardly furnish gains as great as those derived from the first road. Or additional productive agents could be used to improve the quality of the first road. By some method of hard-surfacing, it could be made smoother, more durable, and less subject to the influences of changing weather or seasons. However, the gains derived from getting a good road when a fair one is already in existence would not compare favorably with the gains secured from the building of a fair road when no road at all was in existence. This point suggests that, if enterprisers are allowed to compete for capital goods, an insufficiency of capital in an economic system is likely to result in all business enterprisers having capital goods of fair or low quality rather than in some enterprisers having capital goods of high quality while others have none at all. In other words, enterprisers with no capital goods could afford to offer higher prices for some capital goods, even though the goods were only of fair quality, than enterprisers with fair capital goods could bid for capital goods of high quality.

Again, productive agents could be used to produce additional kinds of capital goods to be used along with the first road. The shoulders of the road could be made broad and firm, or trees and shrubs could be set out beside the road, but once more the gains resulting from such uses of land and labor would hardly be expected to compare with those derived from the first road itself. But, it may be objected, could not the additional agents of production be devoted to producing capital for use in connection with some entirely different kind of production? This, of course, would be

possible but, if we have been wise enough to produce capital goods first for use in those fields of production which are best suited to the application of the roundabout process and in which the gains from the use of capital in production will be greatest, we cannot long avoid by this method the tendency of additional amounts of capital to produce diminishing added gains. All this does not necessarily imply, however, that the diminished gains resulting from the extension of the roundabout process may not be entirely worth obtaining.

This discussion assumes, as it must, that no important advances in technology or the arts of production are made. Under some conditions, the discovery of new and efficient ways of using capital may outrun for a long period of time the increased willingness of people to devote productive agents to the securing of relatively remote rather than immediate results. The discovery of ways of using indirect methods in fields of production in which these methods were once considered inapplicable may permit us to increase our total stock of capital, at least for a time, without any tendency on the part of the additional amounts of capital to furnish diminished gains in comparison with those derived from the original amounts of capital. However, these considerations do not affect our previous conclusion that the extension of roundabout production is likely to produce diminished added gains, for this conclusion deals with a given state of the arts of production.

Indirect Production and Economic Systems. The superiority of indirect production over direct is so generally recognized that we may say that the use of the roundabout process is independent of the type of economic system which is being operated or planned. There may be differences of opinion as to who should own the capital goods, as to the distribution of the gains derived from their use, or as to the basis on which decisions should be reached concerning the total quantity of capital goods to be created or the allocation of these goods among industries, but the use of large quantities of capital in production is considered just as desirable in one economic system as in another. Socialists and communists may not approve of capitalists. They may detest the private ownership of capital and the receipt of interest by private individuals, but the desire to obtain the benefits of using large amounts of capital in production is well-nigh universal.

SPECIALIZATION IN PRODUCTION

Modern production is specialized as well as roundabout. The individual or family does not usually attempt to produce all, or even many, of the commodities and services which are required for want satisfaction. Under pioneering or frontier conditions it has sometimes been necessary for the individual or family to achieve a considerable degree of economic self-

sufficiency, but the average individual of today concentrates his productive energies in a single line of production, or indeed in a single small part of a line of production, and through the process of exchange obtains the commodities and services for his own consumption as these products are turned out by other specialists. And what is true of the individual or family in this connection is also true to a greater or lesser extent of towns and cities, states, regions, and even countries.

Specialization and Other Economic Practices. Specialization, or the concentration of effort upon a restricted phase of economic activity, is closely related to roundabout production, even though the existence of the one practice does not necessarily imply the existence of the other. An individual might make use of a number of different capital goods in production even though he produced several economic goods or performed all the tasks necessary to the production of a single good. On the other hand, individuals might specialize in a given trade or profession, or even in one phase or part of a certain trade, without using any great quantities of capital in the process. In the usual course of events, however, specialization and the roundabout process go hand in hand.

Specialization on the part of the individual worker is also closely related to the development of large-scale production (a matter which we shall investigate in Chapter VII). The manufacture of shoes, for example, could not be divided up into a hundred tasks, with a separate worker assigned to each task, by a firm which produced only three thousand pairs of shoes per year, for in that case each worker would have a chance to perform his task only thirty times per year. Finally, the development of specialization and the development of the exchange process are interdependent. An individual could not specialize in a single trade or task unless he could depend upon being able to exchange his specialized product directly or indirectly for the other commodities and services which he required. Conversely, if each individual or family attempted to be relatively self-sufficient in production instead of specializing, the need for exchanges of products would be minimized.

Advantages of Specialization. The almost universal adoption of specialization in economic activity has been due to the great advantages which specialized production offers in comparison with nonspecialized production. The individual who attempted to carry on many different types of production, or even all the tasks necessary to the production of a single economic good, would find it difficult to reach a high level of efficiency in any one trade or task. He would find himself producing some things or performing some tasks for which his personal abilities were very poorly suited, and he would discover that there were many things which he could not produce for himself at all. Specialization is said to be advantageous because it permits the production of greater quantities of economic goods with a given

expenditure of the agents of production (and hence lowers cost), because it makes possible the production of goods of superior quality, and because it affords us a greater variety of economic goods for our consumption than would otherwise be possible. If we consider it desirable to maximize production and to make the scarce means of production go as far as possible in attaining the ends which we are seeking in our economic activities, then specialization must be regarded as advantageous.

Specialization by Trades. The most obvious type of specialization is that which is ordinarily called specialization by trades. This type of specialization means merely that a man spends all his working time as a farmer, tailor, blacksmith, shoemaker, doctor, lawyer, or something else, instead of dividing his time among a variety of trades or professions, and it tends to occur at an early date in the development of any economic system. By following a single trade or profession all the time, the individual can achieve a greater degree of efficiency in it than would be possible if he had to combine it with several other occupations. He can also select some trade which is suited to his personal characteristics and abilities, whereas, if he followed several trades, he would be almost certain to engage in some lines of production for which he was poorly or indifferently equipped. Finally, he is able to avoid the loss of time which would otherwise occur in changing from one occupation to another. He does not waste time in putting away one set of tools and equipment and getting out another set, and he can keep his capital goods in constant use, instead of having many of them always in idleness as would be inevitable if he followed several occupations which required different kinds of capital goods. Thus it is easy to see why specialization by trades is able to contribute to the advantages of specialization which were mentioned above.

Specialization by Tasks. However, a still greater degree of specialization ordinarily exists in modern production. In most productive establishments today the individual worker does not carry out all or even most of the steps which are required for the production of a finished article. The entire process of production is broken down into numerous small tasks or functions, and separate workers are assigned to the performance of each. It is this specialization by tasks to which the term "division of labor" is most accurately applied. Some years ago students of industry were being told that the making of a ready-made coat had been divided into 39 distinct tasks, that shoe manufacturing included over 100 operations, and that the slaughtering and packing of cattle contained some 230 tasks, while the manufacture of a high-grade watch involved 1,088 separate operations.[2] However, our rapidly changing methods of production soon render such statistics obsolete.

[2] Dexter S. Kimball, *Industrial Economics*. New York: The McGraw-Hill Book Company, Inc., 1929, p. 73.

One significant example of specialization by tasks is found in the assembly line of the modern automobile factory. Such an assembly line, we are told, "consists essentially of a benchlike framework 500 or 600 feet long fitted with guiding surfaces along which the automobile progresses during assembly. At the starting end, the frame of the automobile is placed upon the guides and fastened by a carrier to an endless chain that moves along between the guiding surfaces at the rate of 2 or 3 feet per minute. The several parts of the automobile are supplied by carriers to fixed points on the assembly line in the order in which they are to be assembled. At each station are groups of workers whose sole duty it is to assemble the part that is supplied to them in *the time allowed by the speed of the chain.* Thus, a man may do nothing but insert a given bolt and screw it into place. He may be provided with a wrench that automatically regulates the degree of tightness with which the bolt is set up, thus further reducing the mental and physical requirements of the operation so that almost anyone can be taught to do the work in a short time. It should be noted, also, that the time element is controlled by the chain, so that if a man is compelled for any reason whatever to drop out another must take his place no matter how small or detailed the operation may be, otherwise the entire assembly is held up. By the time the machine reaches the end of the line, it is completely assembled and leaves the line under its own power. Similar methods are employed in assembling automobile engines, which are placed in the automobile during its assembly as a self-contained unit. The rate at which both engine and car can be assembled by these methods is surprisingly great compared to the older methods where the same group of men completed the assembly. The reduction in cost is equally great." [3] Needless to say, in view of the methods employed, the assembly of automobiles and automobile engines requires a small army of specialized workers.

Why Specialization by Tasks Is Advantageous. There are several reasons why specialization by tasks is able to contribute to the general advantages of specialization. If specializing in a single trade, such as tailoring or shoemaking, leads to increased efficiency on the part of the individual workers, we may expect still greater efficiency to result from the constant repetition by the worker of a single simple task in one of these processes of manufacture. The old adage that "practice makes perfect" is doubtless an overstatement, as almost anyone who has taken up golf should be able to testify, but the individual worker, with constant practice, should become quite skillful at a simple industrial task, such as that of attaching a certain kind of bolt to the chassis of successive automobiles on the assembly line. In the second place, the worker should be able to find a task which is well suited to his individual abilities and aptitudes. It is most unlikely that any

[3] Dexter S. Kimball and Dexter S. Kimball, Jr., *Principles of Industrial Organization.* New York: The McGraw-Hill Book Company, Inc., 1939, p. 79.

individual worker will be equally well fitted for all the tasks involved in, say, shoe manufacture, even though that is the best occupation in general for him. Under the system of specialization by tasks, if he cannot become a successful cutter of upper leather, he may be able to get along well in operating a "lasting machine" or in rolling sole leather.

Again, specialization by tasks avoids the loss of time and efficiency which would result from a constant changing by the worker from one task to another. Tools and machines are kept in constant operation and the worker is spared the necessity of putting away and getting out successive types of equipment. It is also well known that most workers do not operate at top efficiency in each working day until they have been at work for a time and have gone through what is sometimes called a "warm-up period." A constant changing of tasks during the day would mean that the worker would seldom, if ever, be working at top efficiency. In the fourth place, under specialization by tasks, many factory jobs for ordinary workers are easy to learn and it is not difficult to transfer from one job to another. A worker today does not usually have to go through an apprenticeship of several years before he can become a full-fledged workman. Indeed, some of our industrialists say that a worker is probably too incompetent to keep on the payroll if he cannot learn to perform almost any job in the factory satisfactorily in a period from a few days up to a week or two. Ordinarily, then, if a worker loses a particular job and business conditions in general are not unduly depressed, he can find and adapt himself to another specialized task quite quickly.

Finally, specialization by tasks tends to stimulate invention and leads to the increased use of machinery. There is little opportunity for the application of mechanical power in manufacturing until this type of specialization has developed. No one would ever have invented a machine which would take in leather as a raw material and produce shoes as a finished product, but when the business of manufacturing shoes was split into numerous small tasks, with many of them requiring only a simple manual operation on the part of the worker, someone was fairly sure to come along with mechanical devices which would perform the tasks more efficiently and economically than they could be carried on by hand. Thus, under the system of division of labor, workers tend to be equipped with large quantities of modern machinery. In some cases, however, the development of machinery has apparently reversed to some extent the trend toward increased division of labor. In some factories, machines are in use which take in a piece of material and perform several operations on it before setting it free for further processing. For example, we are told that "in the plant of a certain large watch manufacturer, . . . a machine has been designed and built at a reported cost of over a quarter million dollars which performs over a hundred distinct operations and has supplanted several

scores of workers."[4] The remaining workers in such cases are still specialized, of course, but the number of specialized tasks performed by individual workers is reduced. This development may therefore be referred to as the integration of tasks.

Other Forms of the Division of Labor. Specialization that is essentially similar to the division of labor among workers in a factory occurs also in the management of our larger enterprises. Sometimes such an enterprise will be divided into departments and a separate executive or representative of management will be at the head of each. Such executives have charge of all the various managerial functions in their respective departments. In other cases the work of management is divided according to function. One man will have charge of purchasing for all the departments of the business, another will control the planning of production in all departments, a third will deal with labor control in all departments, and so on.

The professions also show specialization today that is not very different from specialization by tasks among ordinary workers. Lawyers, instead of trying to cover the entire field of the law, specialize in patent, criminal, corporation, or constitutional law. In medicine, the general practitioner who treated all the ills to which man is heir is now to some extent replaced by the surgeon, the obstetrician, the pediatrician, and so on. Years ago extraction was about the only relief for toothache, and the individual with an aching tooth went either to an ordinary doctor or to his barber for this minor operation. Today dentistry is a separate branch of medicine and in some of the larger centers of population it is further divided so that one dentist specializes in the cleaning of the teeth and massage of the gums (oral prophylaxis), another specializes in the extraction of teeth (exodontia), and so on. Much the same thing is true in the field of education. In our schools and universities at the present time, teachers tend to specialize in particular branches of educational work, as in mathematics, chemistry, romance languages, or economics, and a large university will have a considerable number of people on the staff who teach each of the major subjects. But this is not all. Within the department of economics (or any other department) there is further specialization. Some men teach largely in economic theory, or the history of economic thought, while others are specialists in international economics, economic history, public finance, labor economics, or transportation.

Specialization by Establishments or Enterprises. We have thus far been describing specialization from the point of view of the worker and his job, but in many cases the extent of the division of labor within the factory is closely connected with the degree of specialization characteristic of the enterprise itself. In other words, in the automobile factory only auto-

[4] W. N. Mitchell, *The Organization and Management of Production.* New York: The McGraw-Hill Book Company, Inc., 1939, p. 40.

mobiles are produced, but the factory is frequently merely an assembly point for the products of other specialized concerns which are necessary to the final product. The shoe factory, too, produces only shoes, usually only men's, women's, or children's shoes, in only one or two grades, and in a limited range of sizes, widths, styles, and colors. A paper mill may specialize in producing book paper or newsprint, a textile mill may confine its output to one or two varieties of cloth, or a furniture factory may concentrate on dining room or bedroom furniture.

Specialization by establishments or enterprises has often been accompanied by a high degree of standardization and simplification of products. Each unit of a given type of good is made identical with every other unit of the same good so that parts may be freely interchanged between units, and the number of varieties of each product has been greatly reduced. For example, we are told that the number of varieties of face brick has been reduced from 75 to 2; of common brick from 44 to 1; of paint and varnish brushes from 480 to 143; of metal lath from 125 to 29; of woven wire fencing from 552 to 62; of hotel chinaware from 700 to 279; of milk bottles from 49 to 4; and of beds, springs, and mattresses from 78 to 4.[5] Manufacturers of farm machinery make less than 40 varieties of drills and seeders instead of the former 240, and 30 varieties of ordinary plows instead of 209. Altogether, 227 varieties of farm implements now do the work formerly done by 2,135.[6] An examination of the nature of these goods indicates, however, that there are obvious limits to the application of simplification. It cannot be applied very thoroughly in connection with style goods and novelties where variety is attractive to the user, or in connection with articles which are demanded by different consumers at a considerable range of prices, or where the number of uses to which an article may be put requires large variation in certain characteristics if the different consumers are to be satisfied.[7]

Several reasons may be advanced for specialization by establishments. Concentration on a single product or class of products often makes possible a volume of sales much greater than that which could be realized if the production of the chosen good had to be combined with that of several other articles. Such an increased volume of sales is likely to permit a greater degree of specialization within the plant and the use of the best machines and appliances. Difficulties of supervision are reduced, accounting is simplified, and the planning and control of stocks are facilitated. Thus the net advantages are likely to be found in better goods at lower cost.

[5] A. G. Anderson, M. J. Mandeville, and J. M. Anderson, *Industrial Management.* New York: Ronald Press Company, 1942, p. 286.

[6] *Ibid.*, p. 283.

[7] F. E. Folts, *Introduction to Industrial Management.* New York: The McGraw-Hill Book Company, Inc., 1938, p. 66.

Specialization by establishments is by no means confined to manufacturing. In the field of retail merchandising, the general store, which sold a great variety of articles and cared for practically all of the retail needs of its customers, was once supreme. As rural areas and villages were converted into towns and cities, its place was largely taken by the specialty store which sold only one class of merchandise, such as hardware, groceries, drugs, or dry goods. In some cases a further development has taken place and we find stores which, besides specializing in shoes, sell only men's shoes, women's shoes, or children's shoes. The modern department stores and mail-order houses should not be considered a step backward in the direction of the old general store, for they really consist of a group of specialty stores operating under one management. Wholesalers, jobbers, and other middlemen also often specialize in single classes of products.

The process of specialization by establishments has, however, been reversed in some cases. Many drugstores now handle so wide a variety of articles, such as toilet articles, tobacco products, sporting goods, food, home electrical equipment, and photographic services, that they remind one of the general store. Manufacturing plants which have developed a great volume of output sometimes find that they can now produce more economically articles which they once purchased from other specialized producers. Some plants operate their own foundries instead of buying castings from independent foundries, and large electrical manufacturing companies produce instead of purchasing from other companies such accessories as porcelain, oilcloth, and mica board. Again, electric generators and the steam engines which drove them were once produced in separate establishments, but today the manufacturers of electrical machinery are said to control practically all of the production of turbogenerators.[8] Finally, manufacturers of individual finished products or classes of products sometimes add other finished products to their list. This may happen because it proves impossible to build up an adequate volume of business for the original product or products, because the additional product will permit the company as a whole to have greater seasonal or cyclical stability of operation, or because of other factors.

Geographical Specialization. A final and very important type of specialization is that which occurs in various geographical units. Many towns and cities, for instance, are noted for certain products. Shoe production in the United States centered for many years in the Massachusetts cities of Lynn, Brockton, and Haverhill, with further specialization among them so that women's shoes were made in Lynn, men's shoes in Brockton, and slippers and soft-soled shoes in Haverhill. Troy is well known for its shirts and collars, Gloversville and Johnstown, New York, for gloves, Paterson for its silk mills, Gary, Pittsburgh, and Birmingham for iron and

[8] D. S. Kimball and D. S. Kimball, Jr., *Principles of Industrial Organization*, p. 69.

steel products, Detroit for automobiles, and Los Angeles for motion pictures—these are only a few examples. The concentration of certain industries in certain cities has been accounted for in terms of various combinations of a number of factors: the existence of a supply of power, nearness to markets or raw materials, the presence of a supply of suitable labor or of capital available for investment, favorable climatic conditions, or the advantage of an early start.

States are also noted for producing or manufacturing certain articles, as is Pennsylvania for coke, and for iron and steel products; Connecticut for brassware, copperware, and clocks; California and Florida for citrus fruits; Montana for copper; and Washington and Oregon for fruit and lumber. The wheat, corn, cotton, and other "belts" in the United States are typical of specialization in geographical units larger than a single state. Finally, many countries are well known for their specialization in the production of certain commodities. Examples are the production of silk in Japan, coffee in Brazil, and sugar in Cuba.

Geographical specialization, like the other types of specialization, enables the people of a given geographical region to obtain better and cheaper goods, and usually a greater variety of goods, than could be obtained without it. Nature has scattered her favors widely over the earth, and no one geographical unit can possibly be the best place in the world in which to produce all types of economic goods. But even if one such unit could have advantages over other geographical units in all fields of production, it would be almost certain to have greater advantages in some types of production than in others and a basis for specialization would still exist. Thus, apart from the fact that it may not be possible to produce some things at all in a certain geographical region, we see that any region must be at either an absolute or a relative disadvantage in certain lines of production, and will gain by specializing in other lines and exchanging with other regions. Economic self-sufficiency in a geographical region can be gained only at the cost of having a lower standard of living than would be available on the basis of specialization and exchange.

Specialization and Economic Systems. The study of specialization should not give the impression that this method of organizing production is purely a development of modern times. In fact, specialization by trades, by productive establishments, by geographical regions, and even to some extent by tasks, was practiced in the days of ancient Greece and Rome. The principal factor which limits specialization of most types is the size of the population, and hence of the market which can be served by a certain individual, firm, or industry. Specialization by establishments, trades, and especially by tasks cannot develop fully until the market is large enough to permit the workers and plants to keep busy in their specialized fields of work.

It is also true that the application of specialization is not dependent upon the type of economic system which is in operation. The merits of specialized production are so well understood that specialization by trades, tasks, and productive establishments may be found in Soviet Russia quite as commonly as in capitalistic England and the United States. Specialization was also widely practiced in the late departed fascist systems of Italy and Germany. Specialization by geographical units *within the national boundaries* is also practiced by all types of economic systems. Geographical specialization and exchange between countries have not developed as fully as the other uses of specialization. In the matter of striving for economic self-sufficiency, the fascist nations were the worst offenders, but all of the various types of economic systems have been inclined to ignore the benefits of geographical specialization in the international sphere.

Specialization and Unemployment. Specialization also has several disadvantages. In some cases, these disadvantages are factors which keep specialization from maximizing production to the extent previously implied. In other cases, they are factors which cause specialization to produce allegedly adverse effects on the individual even though it may be successful in increasing production greatly. Under specialization by tasks, the workers are ordinarily gathered together in the factory and operate highly specialized and expensive machinery and equipment; hence it is generally impossible for the workers to own the capital and land necessary to make their labor effective. In the old days, when a worker made a given product from start to finish with the help of productive agents which he himself could own, the problem of unemployment could never become very troublesome. At present, for a chance to work at all the workers are dependent upon those who own or control adequate amounts of land and capital. This means that, if the employment of labor appears likely to be profitable to those who own or control the other agents, the workers have employment; under the opposite conditions the workers are deprived of their jobs and are forced to fall back on savings, charity, or relief, for a time at least. Thus unemployment, largely a problem of modern times and methods, has become one of our most serious economic problems and interferes seriously at times with our ability to make the scarce means of production go as far as possible in the satisfaction of human wants.

Specialization and Labor Problems. When a single worker performed all the tasks of production from raw material to finished product, whether in his own workshop or as an employee of someone else, it was possible to determine with some accuracy just what his labor was worth. Today the worker handles only a single small task in the production of an article, and it is much more difficult to appraise the value of his services. For example, an employee works only on the eyelets of successive pairs of shoes which sell for ten dollars a pair. What is his labor worth? No one really

knows. The employer guesses as nearly as he can, but the results are never entirely satisfactory. The employee is inclined to feel that his wages are too low or that he has to do too much work to get them, while the employer fears that the reverse is true. As a result there is almost constant disagreement, and sometimes open strife, over such matters as hours of work, wages, and other conditions of employment. Industrial conflict interferes with our ability to maximize production. It is difficult to believe that it could plague us so much under nonspecialized production.

Specialization and Monotony. There is another disadvantage of specialized tasks which some people consider very important but which others view rather lightly—the monotony of specialized work. One writer describes it thus: "In one of the great establishments manufacturing automobiles, there is a room filled with punching machines. In front of each machine stands a worker feeding it pieces of steel by hand. A lever is geared to the mechanism, and to this lever the man is chained by a handcuff locked to his wrist. As the punch comes down, the lever moves back taking the hand with it. If for any reason a man wishes to leave the room, all the machinery must be stopped and his handcuffs unlocked by the foreman. To look down the long room is to see machines, levers, and men in unison —feed, punch, jerk back; feed, punch, jerk back. . . . I have heard of no other single task today which so closely approximates the gloomy prophets' picture of the robot—that mechanism of flesh and blood first heard of in a Czechoslovak play, toward which, it is alleged, all men are moving. Yet those workers were handcuffed partly out of sympathy for their flesh and blood. Before the articulated levers were installed, they were, despite the guards, continually losing their fingers and hands under the down thrust of the punch."[9]

Many other harrowing tales are told of the dreadful monotony of specialized tasks. We hear of men in padlock factories who spend all their working days putting the customary two keys on a succession of little brass rings and of girls who sit beside a conveyor belt to earn a living, and a meager one, by thrusting a thumb into countless cardboard boxes to see if they have been properly filled by a machine. Such jobs are supposed to have most distressing effects on these slaves of the machine. Under the weight of utter boredom, workers become careless and are liable to become involved in industrial accidents. Regarding their work as a hateful task to be performed, they have one eye on the clock and the other on their task. Surely such monotonous tasks must be highly depressing, if not maddening, for people who have any intelligence or qualities of initiative and leadership.

There can be no doubt that many specialized tasks are exceedingly

[9] Stuart Chase, *Men and Machines*. New York: The Macmillan Company, 1929, pp. 142–43. Reprinted by permission of the publishers.

monotonous, but there are many offsetting considerations. Although the work is boresome, the working day is relatively short and there are many more attractive ways to spend leisure hours. Many years ago work was less routine and standardized, but the working day often stretched out some fourteen or sixteen hours, and even the most varied work probably became tiresome if continued over too long a period of time. The workers today are also repaid for their monotonous work by having a much larger real income to enjoy than the workers of years ago. It is probable that most workers would prefer a shorter though more monotonous working day, with a larger real income and more leisure time in which to enjoy it, to a longer, more interesting working day with a smaller real income and less leisure time. Specialization by tasks has also taken away from the worker and given to the machine much of the heavy, dangerous, dirty work which was once performed by hand.

It is perhaps true that a person of high intelligence, initiative, and leadership might suffer a nervous breakdown if he did not lose his reason while working unrelentingly at some modern specialized task. But in practice such an individual does not often stay long at these occupations. For many more ordinary individuals such simple routine tasks may be just what the doctor ordered. In a few days the task is thoroughly learned and mental concentration is no longer necessary. The worker can function efficiently while his mind wanders to more attractive fields, and he thinks of the motion picture he saw the night before, of the tennis match he will play after work, or of the trip he will take with his family in the car on Sunday. In fact, employers who have proposed to make the work more interesting for their men, by shifting them about periodically from one job to another, have sometimes been threatened with strikes to prevent the operation of their program. This attitude on the part of workers suggests that the writers who have dwelt so feelingly on the effects of the monotony of specialized tasks have been mentally putting themselves in the workers' shoes. That is, they have described the effects they think such work would have on themselves and not its effects on workers who actually hold these jobs.

Specialization and Loss of Pride in Work. Another alleged drawback of specialization by tasks is that the worker can no longer take pride in his work and in his product. The tailor of many years ago who produced an entire suit of clothes could point with pride to the finished article as distinctively his own. Today the same man may work only on buttonholes in a clothing factory. He can no longer thrill with satisfaction at the completion of a suit—indeed he may never even see the finished product. There is probably some truth in this criticism of specialized production, but it is not strictly necessary that all workers be proud of what they do for a living. They will ordinarily have the time and the facilities for other

activities, outside their working hours, from which they can derive both pride and satisfaction.

Specialization and One-sided Development. Finally, it is sometimes remarked that a worker who spends years at a simple specialized task is likely to be cursed with one-sided development. Certain sets of muscles, used in the daily task, are overemphasized while the development of other muscles and even the development of the worker's mind are neglected. Here again is something which may or may not actually occur under specialization by tasks. Whether or not it happens depends upon how the worker uses the leisure time and the income which specialized production places at his disposal.

Specialization and Interdependence. In turning from specialization by tasks to other phases of specialization, we may note that the extremely specialized plant may suffer serious financial loss if new inventions or other changes in productive methods are made, and an industry which is concentrated in a certain geographical region may be ruined if changing economic conditions make it more advantageous to manufacture its specialty in some other section or area. Under specialization our various industries are so highly interdependent that a severe disturbance in any one field of production is likely to have a far-reaching influence on other parts of the economic structure. A nonspecialized plant or firm which turns out several products may not suffer greatly during a business depression for, while the demands for some products fall off sharply, the demands for others may be well maintained. On the other hand, a specialized establishment or firm with a single product may be hard hit if the demand for this good declines drastically. In similar fashion, a plant with several products may find it possible to stabilize productive operations throughout the year, but a specialized plant may have to run at capacity in one season and at a small fraction of capacity in another part of the year. Thus specialization may be associated with business instability.

Similar results may be traced to geographical specialization. Economic activity in different geographical areas tends to be prosperous or depressed at the same time and the interdependence of geographical areas for various products, especially among countries, may lead to unfortunate results when trade is cut off or disrupted in time of war. On the whole, however, in the opinion of most students of the subject, the various advantages of specialized production seem to outweigh its disadvantages.

QUESTIONS AND PROBLEMS

1. "Production is the creation of commodities which have utility." Discuss.
2. Indicate some of the ways in which utility is created in production.
3. "Very few economic activities can be classified as unproductive unless one shifts from the notion of production as the creation of utility to some

concept of production which requires the enhancement of the welfare of society." Explain.

4. "Competitive advertising is an unproductive activity." Show whether you agree.

5. "The members of the armed forces of the nation during World War II were engaged in unproductive activity." Discuss.

6. "Human exertion which produces money or real income is classed as labor but that which produces only psychic income is not labor." Explain.

7. Distinguish between direct and indirect (or roundabout) production.

8. Show how the production of some commodity may become increasingly roundabout as time goes on.

9. "The nature of the cost of obtaining capital goods is the same in all economic systems." Show whether you agree.

10. State the "benefits" of indirect production and indicate the point of view from which they may be regarded as benefits.

11. "Indirect production appears to yield less the further it is extended." Do you agree? Explain.

12. "The use of indirect production is not at all dependent upon the type of economic system which exists in a particular country." Show whether you agree.

13. Distinguish between specialization by trades and specialization by tasks, and illustrate each of these types of specialization.

14. "It is necessary to distinguish the advantages of specialization by tasks from the reasons why specialization by tasks is able to furnish these advantages." Explain.

15. "Specialization by tasks occurs in the professions and in the management of productive enterprises, as well as in the activities of ordinary workers." Explain.

16. Explain and illustrate specialization by productive establishments or enterprises, and geographical specialization.

17. "Specialization is a development of modern times and is peculiar to capitalistic economic systems." Show whether you agree.

18. "The advantages of specialization by tasks outweigh the disadvantages." Discuss.

19. "While the monotony of work, the one-sided development of workers, and the loss of pride in work are most frequently mentioned as disadvantages of specialization by tasks, there are other disadvantages which are more important than these." Do you agree? Explain.

20. "Specialization by individuals in production is desirable in part because it enables us to take advantage of the varying abilities of these individuals, but such specialization would still be desirable even if all individuals were endowed with precisely the same abilities." Show whether you agree.

REFERENCES FOR FURTHER READING

Black, J. D., *Production Economics.* New York: Henry Holt and Company, 1926, Chaps. 6–10 and 14–16.

Bohm-Bawerk, Eugen v., *The Positive Theory of Capital.* New York: G. E. Stechert and Company, 1923 (reprint), Book I, Chaps. 2–3, and Book II, Chaps. 2–3.

Brown, H. G., *Economic Science and the Common Welfare.* Columbia, Mo., Lucas Brothers, 1936, Part II, Chap. 3.

Clark, J. M., *The Economics of Overhead Costs.* Chicago: University of Chicago Press, 1923, Chaps. 1–3.

Folts, F. E., *Introduction to Industrial Management.* New York: The McGraw-Hill Book Company, Inc., 1938, Chaps. 2–3.

Kimball, D. S., *Principles of Industrial Organization.* New York: The McGraw-Hill Book Company, Inc., 1933, Chap. 5.

Marshall, Alfred, *Principles of Economics.* London: Macmillan & Company, Ltd., 8th ed., 1920, Book IV, Chaps. 9–10.

Taussig, F. W., *Principles of Economics.* New York: The Macmillan Company, 4th ed., 1939, Chap. 2.

VI

Proportions of the Productive Agents

As we have seen, the production of any economic good requires some combination or other of the agents of production. However, though existing technology and methods of production set some limits on the proportions in which the agents of production may be combined, it is usually possible to produce any given economic good by using the agents of production in various proportions or combinations. In agriculture, it is possible to raise a crop by using large quantities of land with relatively little labor and capital to the acre or by using small amounts of land with relatively large quantities of labor and capital to the acre. Even with a given quantity of land, we may carry on cultivation with much labor armed only with simple tools and instruments or with little labor assisted by a great deal of expensive and complicated equipment. If a given amount of earth is to be excavated, we may use a large number of workers with shovels and wheelbarrows for equipment or a much smaller number of workers assisted by steam shovels and dump trucks. Similarly, in industry, we may manufacture shoes by using many workers who have only small quantities of capital goods or a smaller number of workers with relatively large quantities of capital per worker.

Even in cases in which it would seem that the proportions of the productive agents were closely fixed by the methods of production used, there may actually be some possibilities of variation. In many factories, for example, it would appear that one, and only one, worker would be required for each machine, but in fact additional workers could be used in supervising the work of the machine operators, in keeping the machines and equipment always in the best possible condition, and in testing and checking the final product. Thus enterprisers usually have a choice as to the proportions in which they will combine the agents of production, and it is their task to make such choices. In the present chapter, we shall examine some of the results of combining the productive agents in varying proportions and we shall find that the efficiency of production and the resulting costs of the final products are significantly affected by the proportions in which the productive agents are combined.

THE LAW OF DIMINISHING PRODUCTIVITY

Agricultural Production with Labor Variable. In the first place, let us consider the results which might be obtained in agricultural production with labor as the variable productive agent. Specifically, let us say that we have one unit of land (possibly an acre) and given amounts of seed, tools, and equipment, which we may use in conjunction with varying numbers of units of labor (with each unit of labor consisting of two days of labor of given quality spread over the growing season). If we used only one unit of labor with the fixed amount of land and equipment, our cultivation would be extensive in character. The labor would be spread thinly among the tasks necessary to cultivation and each task would have to be performed rather hurriedly and superficially. As a result, we could not expect a large output of the desired crop. As we applied more and more units of labor in conjunction with the fixed productive agents, cultivation would become more intensive. The soil could be prepared for cultivation more carefully, we could take better care of the growing crop, and each task could be performed more thoroughly. The result would be an increasing yield of product from our plot of land. However, the yield could not be made to increase indefinitely. As we went on adding units of labor, we should find that it became progressively more difficult to increase the crop, and sooner or later we should reach a point beyond which further units of labor would add nothing to the total crop and might even decrease it.

The possible results of our experiment with various proportions of the productive agents may be seen more exactly in Table 1. The first two columns of this table speak for themselves. They show the varying number of units of labor which we might use with our fixed land and equipment, and the total yield or output of product which might be associated with the use of these varying quantities of labor. In the third column, labeled marginal output, we show the amount of product added to the total by each successive unit of labor. For example, when we add the sixth unit of labor, we cause the total output to increase from 51 bushels to 60; therefore we say that the marginal or added output at this stage is 9 bushels. Another way of looking at the marginal output is to say that it is the amount of product which would be lost at any given stage of cultivation if one unit of the variable factor (labor) were withdrawn from use. When four units of labor are employed, the withdrawal of one unit of labor would cause the total product to decline from 40 bushels to 28. Hence, the marginal product of labor is 12 bushels when four units of labor are in use. Average output per unit of labor, as shown in the fourth column, is found simply by dividing each successive total product by the number of units of labor used to produce it. For example, a total output of 75 bushels when

nine units of labor are used gives an average product of 8.33 bushels per unit of labor.

Table 1: Results of Agricultural Production with Labor as the Variable Productive Agent

Number of Units of Labor Employed	Total Outputs (*in bushels*)	Marginal (Added) Outputs (*in bushels*)	Average Outputs per Unit of Labor (*in bushels*)
1	6	6	6.00
2	16	10	8.00
3	28	12	9.33
4	40	12	10.00
5	51	11	10.20
6	60	9	10.00
7	67	7	9.57
8	72	5	9.00
9	75	3	8.33
10	76	1	7.60
11	76	0	6.91

The data in Table 1 indicate that the total product would go on increasing as we applied from 1 to 10 units of labor with our fixed land and equipment, but that the eleventh unit of labor would produce no further increase. Even while the total output is increasing, however, we note that considerable changes take place in the marginal product of labor and in the average output per unit of labor. The marginal product is only 6 bushels when one unit of labor is used. As we increase the number of units of labor used and make the amount of labor employed more adequate with respect to the fixed productive agents, the marginal output increases to 10 bushels and then to 12. After remaining at 12 bushels when the fourth unit of labor is applied, the marginal output gradually decreases as the quantity of labor used becomes oversufficient with respect to the fixed agents, and reaches zero when 11 units of labor are used. While all this is going on, the average output per unit of labor increases from 6 bushels to a maximum of 10.2 bushels and then declines to 6.91 bushels when 11 units of labor are used.

Thus there is a tendency for both the marginal product of labor and the average product per unit of labor to diminish, after a certain point has been reached, as varying numbers of units of labor are used in conjunction with a fixed quantity of land and equipment. There remains the task of deciding exactly where we should say the diminishing productivity of labor sets in. Since the average product per unit of labor goes on increasing even after the marginal output of labor begins to decline, and since we are not getting the maximum results from our use of labor until

the average product per unit of labor reaches its peak, we define the point of diminishing productivity in terms of the average product per unit of labor. Thus the point of diminishing productivity is said to be *the point after which the average product per unit of labor (or any other variable productive agent) starts to decline.* In Table 1, this point is reached with the application of the fifth unit of labor, since the use of the sixth unit would cause the average product to decrease from 10.2 bushels to 10.

Curves of Average and Marginal Productivity. The relationship between the marginal product of labor and average product per unit of labor can be seen clearly in a diagram, such as Figure 1. In this diagram, the units of labor from one to eleven are laid off along the horizontal axis *(OX),* with each space representing one unit of labor. On the vertical axis *(OY),* we measure off bushels of product, with each space representing one bushel of product. From Table 1 we know that the average product per unit of labor is 6 bushels when one unit of labor is used. This situation is represented in Figure 1 by the point where a horizontal line drawn to the right from the *OY* axis at the level of 6 bushels would intersect a line drawn perpendicularly upward from the point on the *OX* axis which represents one unit of labor. The other marginal and average products of labor are plotted in similar fashion. The average product curve (*AP*) is a line drawn to connect the various points which represent the average products of labor associated with varying numbers of units of labor, and the marginal product curve *(MP)* is similarly a line which connects the points which represent the marginal products of labor as various quantities of labor are used.

In Figure 1 we note that the marginal product curve rises more rapidly than the average product curve, reaches its peak at an earlier stage in the application of successive units of labor to the fixed productive agents, and then declines more rapidly than the average product curve. It is also apparent that the marginal product curve intersects the average product curve at the highest point of the latter curve. These relationships are readily explained. As long as the marginal product added by a given unit of labor is greater than the average product of all preceding units of labor, the average product per unit of labor must rise. Thus, in our illustration, the fourth unit of labor has a marginal product of 12 bushels while the average product of the first three units of labor is only 9.33 bushels, so that the average product rises to 10 bushels when four units of labor are used. In this connection it makes no difference whether the marginal product itself is rising or falling. Just so long as it is above the average product, the average product must rise. On the other hand, when the marginal product resulting from the use of a given unit of labor is smaller than the average product of all preceding units of labor, the average product of labor must fall. In our illustration the eighth unit of labor has a marginal product of

5 bushels while the average product of 7 units of labor was 9.57 bushels, so that the average product declines to 9 bushels when 8 units of labor are used. Since marginal product is greater than average product when average product is rising, and smaller than average product when average product is falling, it follows that marginal product will be equal to average product when average product is neither rising nor falling; i.e., when aver-

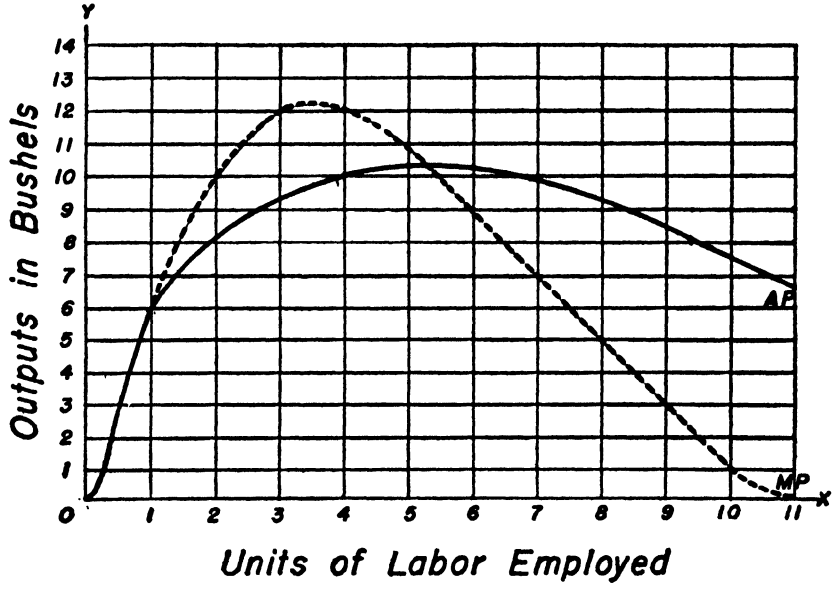

FIGURE 1.—Curves of Average and Marginal Product

age product is at its maximum. This means also that marginal product is equal to average product at the point of diminishing productivity.[1]

Diminishing Productivity and Practical Cultivation. Our illustration has shown the productive results which might be obtained by using all the way from one to eleven units of labor in conjunction with a fixed amount of land and equipment. In practice, however, if the cultivator is interested in obtaining the best possible financial results, actual cultivation would fall somewhere in between the point where the average product per unit of labor reaches its maximum (point of diminishing productivity) and the point where the marginal product of labor reaches zero (point of absolutely diminishing returns). Let us see why this is true. If production were carried beyond the point where marginal product becomes zero, marginal

[1] In Figure 1, it is apparent that the point of diminishing productivity is reached when slightly over 5 units of labor are employed. However, in terms of the whole units of labor used in Table 1, it may still be said that the point of diminishing productivity is reached when 5 units of labor are employed.

product would be negative, average product per unit of labor would continue to decline, and the total product would also decline. Clearly there would be no excuse for using more of the variable agent (labor) to obtain a smaller total product. Even if labor were costless, there would be no reason to use it beyond the point of absolutely diminishing returns.

If we carried production back beyond the point of diminishing productivity, we should be increasing the proportion of the fixed productive agents (land and equipment) to the variable productive agent (labor) while accomplishing only a decrease in the average product per unit of labor and also a decrease in the total product (or product per unit of the fixed productive agents). Obviously there would be no reason to desire this result. To put the matter another way, until the point of diminishing productivity is reached, we are getting increasing product per unit of the variable factor and also increasing product per unit of the fixed agents (total product), and we should never stop production so long as this tendency continues. Even if the fixed productive factors were costless, it would always pay to apply the variable factor to them until maximum output per unit of the variable factor had been reached.

Our conclusion is that the only economically significant sector of our illustration is that area in which average product per unit of the variable agent is declining while product per unit of the fixed agents (total product) is increasing. This area, of course, lies between the point of diminishing productivity and the point of absolutely diminishing returns. Just where, within this area, cultivation should be stopped will depend upon the price of the product and the cost of the fixed and variable agents of production. We shall examine this matter later in the present chapter.

Agricultural Production with Land Variable. So far we have been dealing with agricultural production using labor as the variable productive agent, but actually any productive agents may be fixed and any agent may be variable. If we used fixed quantities of labor, tools, equipment, and seed in conjunction with variable quantities of land, for example, the productive results would be similar to those which we have already discussed. Let us suppose that we have one hundred units of labor at our disposal along with equivalent quantities of various necessary capital goods. If we piled all these fixed productive agents on a single acre of land, the land would be overcultivated and the productive results would be relatively unsatisfactory. As we spread the fixed factors over additional units of land, the quantity of land would become relatively more sufficient with respect to the fixed agents, and we should get an increasing product both per unit of the variable agent (land) and per unit of the fixed agents (labor and capital).

If we continued to add units of land, however, without increasing the quantity of labor and capital at our disposal, the quantity of land would

become oversufficient with respect to the fixed labor and capital and each unit of land would be somewhat undercultivated. The result would be a declining average product per unit of land, though the total product, or product per unit of labor and capital, would continue to increase. Finally, we should reach the point where an attempt to spread our fixed labor and capital over more units of land would add nothing to the total product which we could obtain or would even decrease it. Tables and diagrams comparable to Table 1 and Figure 1 could be constructed for agricultural production with land as the variable agent, but they would add little to what has already been presented.

Industrial Production with Labor Variable. Having seen that the results obtained by using varying proportions of the productive agents are the same regardless of the identity of the fixed and variable agents, we should now note that these results are not confined to agriculture but may be obtained in any field of economic activity. Let us suppose that we have a shoe factory, located on a suitable piece of land, and equipped with many and varied types of shoe machinery, tools, a power plant, storage and packing rooms, and adequate amounts of raw materials, supplies, and fuel. This enterprise could turn out some product with only a few workers, just so there were enough to generate the necessary power, operate at least one unit of each type of machinery part time, and perform other necessary tasks. However, much of the power which was generated would be wasted, much of the machinery and other facilities would be idle most of the time, and each worker would have to be a kind of one-man band or Jack-of-all-trades. The total product and output per worker would be rather small.

As more workers were used along with the fixed productive agents, the quantity of labor would become relatively more sufficient, better use could be made of the fixed productive facilities, and the workers could concentrate on individual specialized tasks. The result would be both an increasing total product and an increasing average product per unit of labor at this stage of operation. Sooner or later, however, the tendency to diminishing productivity per unit of the variable productive agent would manifest itself. There are obvious limits on the specialization of labor by tasks, as we have seen, and after a certain point had been reached additional workers would have to be assigned to less and less important tasks in the enterprise. In this stage each worker added would still be able to increase the total product somewhat, but the average product per unit of labor would decline as additional units of labor were employed. Finally, the point of absolutely diminishing returns, beyond which each additional worker would add nothing to the total product or would even cause it to decrease, would be reached.

Other Illustrations. Similar results would be obtained in using variable proportions of the productive agents in other fields—for example, in adding

successive clerks in a store of given floor space, equipment, and stock in trade, or in sending more and more miners into a coal mine with given working space and equipment. Moreover, we should experience the same general results in constructing a building on a given site. After a certain number of units of capital and labor had been applied on the fixed amount of land and the building had reached a certain height, we should find that each successive unit of labor and capital would add a smaller amount to the total of floor space available for use and that the average product per unit of labor and capital would decline. Of course, in the case of a building, we are likely to decide in advance on the size which each floor should have and then proceed to carry out our plan even though, beyond a certain point, it becomes necessary to add more and more units of labor and capital in order to create another floor of given size. This practice may conceal the *point* of diminishing productivity but it does nothing to affect the *tendency* to diminishing productivity. The average product per unit of the variable productive agents declines whether we add a larger unit of the variable agents to get a given addition to the total product or add a given unit of the variable agents to get a smaller addition to the total product.

The Law of Diminishing Productivity. Now that we have observed something of the universality of the results obtained by combining the productive agents in varying proportions we are in a position to state the general law—known as the Law of Diminishing Productivity—which covers all such situations. This law states that, *if the technology of production remains unchanged and equal successive units of some one agent of production are applied to a fixed amount of other productive agents, the average product per unit of the variable agent, after a certain point has been reached, will decline.* The law is, in effect, a statement that the various agents of production are imperfect substitutes for each other. If they were perfect substitutes for each other, more and more units of a variable agent could be used in conjunction with fixed amounts of other agents without encountering any tendency for marginal and average product per unit of the variable agent to decline.

Some of the conditions under which this law operates perhaps deserve additional emphasis. For one thing, the law operates only when successive units of some variable agent of production are used in conjunction with a *fixed amount* of other productive agents. If, for example, when the application of successive units of labor had been carried to the point of diminishing productivity on a given piece of land, additional units of land of equal quality could be readily obtained and the remaining units of labor applied on this new land, a declining average product per unit of labor would never have to be experienced. On the other hand, if all land of the given quality for the purpose at hand were already in use and any new land which could

be brought into use would be markedly inferior in quality, it would no longer be possible to avoid getting a declining average product per unit of labor on the original land by bringing more land into use.

Again, the law presupposes that all the agents of production are of given quality and are combined according to known methods of production. If labor is the variable productive agent, the successive units of labor are taken to be equal in quality as well as in size. The diminishing productivity of labor does not depend on the use of labor of poorer and poorer quality, but rather on the less and less efficient use of labor of given quality. The quality of any factor used in fixed amounts in a given case is also taken to be uniform and given. If land is a fixed factor, the diminishing productivity of labor applied to it is not in any way dependent upon the exhaustion of the fertility of the soil. Diminishing productivity is not the result of any errors in judgment on the part of management. The tendency operates under management of any given quality and state of knowledge. Whether capital is a fixed or variable agent in a given situation, the conditions under which the law operates exclude new inventions. Finally, it is assumed that the productive agents will be used according to some known method of production. All this does not mean that the enterpriser is assumed to have no choice as to the grade of land, labor, or capital goods which he will use or as to the productive methods which will be employed. The enterpriser may use any available grades of any productive agents and any of the several methods which may be available for producing a good, but new inventions and improvements in productive methods are excluded.

The assumption of given sets of productive methods and a given state of technology leads some people to erroneous conclusions concerning the Law of Diminishing Productivity. That is, they regard the law as one which may come into effect at some time in the future when our inventiveness begins to lag, but they think the law has had little significance in the past because of constant improvement in machines and methods of production. Actually the law has always operated in the past and it is operating now. If, at any given time, some units of the variable factor were improved in quality with respect to the other units or were applied to the fixed amounts of other agents in new and improved ways, the functioning of the law would be interfered with, it is true, but *under any given sets of productive methods and appliances* the law operates as stated regardless of how good or poor these methods and appliances may be. The fact that we now have, in most fields of production, better machines and methods of production than formerly does not keep the law from operating in the present state of the arts of production. The operation of the law requires that productive methods and technology be fixed *at any given time* but it does not require that they be fixed *through time.*

Finally, the law states that the average product per unit of the variable agent will decline only "after a certain point has been reached." This qualification is necessary because, starting from a situation in which very little of the variable agents is used, both the average and the marginal product of the variable factor may increase for a time as the quantity of this factor is made relatively more sufficient in relation to the fixed factors. The exact location of the "certain point," or point of diminishing productivity, may vary considerably from one situation to another. Indeed it is possible that, in some cases, the average and marginal product of the variable agent will begin to decline after the first unit of the variable agent has been applied. Hence the law does not state definitely that there will always be an area in which the average product per unit of the variable factor will increase, but merely suggests the possibility.

Diminishing Productivity and Economic Systems. The Law of Diminishing Productivity, as we have seen, has to do with physical quantities of the productive agents (acres of land, days of labor, numbers of tools) and physical quantities of product (bushels of wheat, pairs of shoes). It makes no assumptions concerning the existence or nonexistence of any particular economic institutions, and its operation does not depend upon prices, costs, or any other financial considerations. Consequently, it is the type of law which may be expected to operate quite as well in one type of economic system as in another. It is as significant in Soviet Russia or in a fascist system as it is in the United States. If it is necessary in these other economic systems, as it is in our own, to try to increase production at times by expanding the quantity of some variable agent of production which is used in conjunction with a fixed amount of other productive agents, transferring the ownership of land and capital or the control of economic activities in general from private individuals to society as a whole or to the government will not interfere with the working of the law. The great dictator can no more prevent the operation of the Law of Diminishing Productivity than he can prevent the ebb and flow of the tides or the coming of night after day.

DIMINISHING PRODUCTIVITY AND COSTS OF PRODUCTION

We have seen that the actual rate at which an enterprise will be operated (in combining fixed and variable agents of production) will fall somewhere between the point of diminishing productivity and the point of absolutely diminishing returns, but we do not yet know how its exact rate of operation will be determined. In the present section we shall see that this rate of operation depends on the price which can be obtained for the product and the costs of production which result from and are the counterpart of the productive tendencies which have already been observed. The situation under discussion, in which some agents of production are

want to add units of product to his output until he reaches the point beyond which the additional product resulting from the use of one more unit of the variable factor would add more to his costs than to his income. Let us see how this works out. If the price of his product were $1.60 per bushel, the enterpriser would use 8 units of labor and turn out 72 bushels of product, since marginal cost at this output would be $1.50 but would rise to $2.50 if 9 units of labor were used and total output became 75 bushels. However, at an output of 72 bushels, average cost per bushel would be only $.903 and, by subtracting this average cost from the selling price of $1.60, we find that the enterpriser would be making a profit of $.697 per bushel or $50.18 on 72 bushels. This total profit is greater than that which he could make on any other output when the price of the product is $1.60.

There are two ways of showing that total profit would be maximized at an output of 72 bushels. First, if the enterpriser stopped at an output of 60 or 67 bushels, he would be failing to turn out some units of product which would add more to his income than to his costs. Conversely, if he went on to an output of 75 or 76 bushels, he would be adding to output some units of product which would increase costs more than they would increase income. Since either of these procedures would reduce total profit to some extent, the output which will provide greatest total profits is 72 bushels In the second place, it is easy to show by computation that the total profits would be less at any output larger or smaller than 72 bushels. If only 60 bushels were produced, average cost would be only $.833 per bushel and profit per bushel would be $.767, but total profit would be only $46.02. Similarly, if 67 bushels were produced, total profits would be about $.50 less than at an output of 72 bushels. If 75 bushels were produced, average cost would be $.967, profit per bushel would be $.633, and total profit would be only $47.48. Once again we see that the most profitable output would be 72 bushels when the price of the product is $1.60.

In the case just discussed, the output which provided the greatest total profit was well beyond that which we have described as normal or optimum output. However, under other conditions, exactly the normal output might be produced. If the price of the product were $.833 per bushel, the illustration indicates that the enterpriser could use 6 units of labor and produce exactly the normal output of 60 bushels without having marginal cost rise above price. Since average cost is equal to marginal cost at normal output, it is clear that the enterpriser would just "break even" and would realize no net profit on his operations. However, once again, we should find that the enterpriser could not secure a better financial result by producing a larger or a smaller output.

If the price of the product were only $.75 per bushel, we see from Table 5 that the enterpriser could produce an output of 51 bushels (with 5 units of labor) without having marginal cost rise above price per bushel.

However, at this output, Table 3 informs us that the average cost would be $.833 per bushel, so that productive operations would have to be carried on at a loss. For this reason, some students would conclude that the enterpriser would be better off if he did not operate his enterprise at all under this unfavorable price situation for the product. This conclusion is erroneous because it fails to consider the fact that the enterprise has fixed costs of $5 which would have to be paid even if the enterprise did not operate. By producing 51 bushels and selling them for $.75 per bushel, the enterpriser would cover his average variable costs of $.735 per bushel at that stage (Table 2) and have 1½ cents per bushel left over, or $.765 altogether, to apply to his fixed costs of $5 which would be a total loss if he did not operate. Thus the enterpriser once again secures the best financial result (in this case, the minimum total loss) by producing the output beyond which further units of product would add more to cost than to income, for it can readily be shown that the enterpriser could not reduce his total loss under a price of $.75 per bushel by producing an output larger or smaller than 51 bushels.

Finally, if the price of the product were only $.65 per bushel, the enterpriser would secure the best financial result by not operating at all. To be sure, this price would more than cover the marginal cost of $.625 at outputs of 40 and 28 bushels, but it would not cover the average variable costs of $.75 and $.804, respectively, at these outputs. Since the enterpriser could avoid paying the variable costs altogether by ceasing to operate temporarily, he would have no incentive to operate under a price of $.65 per bushel for his product. Our general conclusion is that, if it pays to operate an enterprise at all under a given price for the product, the output which will provide the best financial result (maximum total profit or minimum total loss) will be that output beyond which further units of product will have a marginal cost higher than the price, or income per unit, which they will furnish the enterprise.

QUESTIONS AND PROBLEMS

1. Show how an enterprise in agriculture, manufacturing, or commerce could operate with varying proportions of the productive agents.

2. Distinguish between total, marginal, and average product.

3. "The point of diminishing productivity is the point beyond which the enterpriser can no longer make a profit." Show whether you agree.

4. "The actual rate of operation of an enterprise, in combining variable and fixed productive agents, will fall somewhere in the range between the point of diminishing productivity and the point of absolutely diminishing returns." Why?

5. "The Law of Diminishing Productivity is a statement of several assumptions or necessary conditions, and one conclusion." Explain.

6. "The Law of Diminishing Productivity states, in effect, that the various

agents of production are imperfect substitutes for each other." Do you agree? Explain.

7. "If labor is the variable productive agent in a given situation, the diminishing productivity of labor depends on the use of labor of poorer and poorer quality." Discuss.

8. "The operation of the Law of Diminishing Productivity depends upon fixed productive methods and technology. If these factors improve through time, the Law has no significance." Show whether you agree.

9. "The Law of Diminishing Productivity has significance only in capitalistic economic systems." Discuss.

10. Distinguish between fixed and variable costs; between average and marginal costs.

11. "Variable costs are those whose total amount varies directly and proportionally with output." Do you agree? Explain.

12. "The point of optimum cost will usually coincide with the point of diminishing productivity." Discuss.

13. What is meant by normal output? Explain.

14. "Marginal cost is equal to average cost when average cost is at a minimum." Why?

15. "The competitive enterpriser will go on adding units of product to his output as long as marginal cost does not exceed the price per unit of product." Do you agree? Explain.

16. "The most profitable output for a competitive enterpriser is that at which price is equal to marginal cost." Discuss.

17. "If the price of a good, though equal to marginal cost, is less than average cost, an enterpriser cannot afford to operate." Show whether you agree.

18. "If the price of a good, though equal to marginal cost, is less than average variable cost, an enterpriser cannot afford to operate." Discuss.

19. "The actual output selected by an enterpriser in a relatively short period of time will always be the output at which marginal cost is a minimum." Show whether you agree.

20. "If the cost of the fixed productive agents increases while that of the variable agent decreases, the optimum combination of the agents will include more units of both fixed and variable agents than formerly." Do you agree? Explain.

REFERENCES FOR FURTHER READING

Black, J. D., *Production Economics.* New York: Henry Holt and Company, 1926, Part III, Chaps. 11–13.

Boulding, K. E., *Economic Analysis.* New York: Harper & Brothers, 1941, Chap. 22.

Carver, T. N., *The Distribution of Wealth.* New York: The Macmillan Company, 1913, Chap. 2.

Marshall, Alfred, *Principles of Economics.* London: Macmillan & Company, Ltd., 8th ed., 1920, Book IV, Chap. 3.

Stigler, G. J., *The Theory of Competitive Price.* New York: The Macmillan Company, 1942, pp. 116–128.

VII

The Size of the Productive Unit

In a relatively short period of time, a firm will have at its disposal fixed amounts of certain productive agents (such as its site, plant, and machinery). In this situation changes in output are accomplished by using varying quantities of other productive agents. We have already examined the results obtained, by combining the productive factors in varying proportions, in terms of productivity, efficiency, and cost. While the total output of the firm will go on increasing through a considerable number of stages of operation, the average product per unit of the variable agent or agents will first increase and then decrease as the quantity of the variable agents in use increases from very small to adequate to very large in relation to the fixed agents. When translated into terms of cost, these variations in output cause the average cost per unit of product first to decrease and then to increase as larger and larger quantities of the variable agents are used with fixed quantities of other agents.

However, in longer periods of time, the firm's desire to increase its output need not depend upon the use of larger quantities of some agents of production in combination with fixed quantities of other agents. That is, it is possible to obtain larger quantities even of those agents of production which were fixed in the short period and to increase output by increasing the total size or scale of the firm or establishment. In the present discussion we are interested in this latter development and must ask what is likely to happen to the firm's efficiency and average cost of production as it increases output by increasing the quantities of all the productive agents which it uses.

SIZE OF THE PRODUCTIVE UNIT IN RELATION TO COST

The General Question of Divisibility. In any industry or field of production we find that the efficiency of the individual firm or productive unit increases and its average cost of production per unit of product decreases as the size of the firm increases from very small to what may be called optimum size. In the preceding chapter we noted that the tendency to dimin-

ishing productivity per unit of a variable productive agent depends upon the fact that the individual agents of production are imperfect substitutes for each other. In the present chapter we note that the tendency to increasing efficiency and decreasing average cost as a firm expands its total size or scale toward the optimum size results from the fact that the individual agents of production are imperfectly divisible. If each productive agent were perfectly divisible, one firm could combine small quantities of the productive agents in a given proportion and obtain just as great efficiency and just as low cost as another firm which combined large quantities of the productive agents in the same proportion. We must now go on to see how this indivisibility of the productive agents manifests itself in connection with the organization and operation of particular firms.

Indivisibilities of Machinery and Plant. The effects of indivisibility are clearly seen in connection with the machinery and plant of a productive enterprise. In some cases, machines or pieces of equipment cannot fall below a certain minimum in size if they are to be useful at all. This would be true, for example, of the conveyor system in an automobile factory. More commonly, however, machines or pieces of equipment can be made in a great variety of sizes. Here indivisibility shows itself in the fact that the increased productive capacity of the larger machine or piece of equipment does not require anything like a proportionate increase in cost, or in the quantity of other productive agents necessary to construct and operate it.[1]

If we want one machine to be able to render twice as much productive service as another which is of relatively small size, it will not require twice as much material in its construction or twice as much labor to produce its parts and put them together. It will not occupy twice as much space, need twice as much power to operate it, or require twice as much labor to tend it and keep it in repair. In similar fashion, if one factory building is to have twice the productive capacity of another which is relatively small, it will not, in a given location and under given methods of construction, have to be twice as large or cover twice as much ground, it will not require twice the amount of construction materials in its walls, roof, and floors, its heating will not require twice as much fuel, and so on. The result of all this is that, as a plant or productive unit increases in size (within limits), the cost of plant, machinery, and equipment per unit of product will decline.[2]

[1] Thus, while the volume or capacity of a ship increases at least roughly as the cube of its dimensions, its water resistance, and hence its need for fuel, equipment, and other things increases far less rapidly. The capacity of a metal boiler also increases as the cube of its dimensions, but the quantity of metal used in its construction would increase only as the square of its dimensions if the thickness of its walls did not have to be increased to some extent.

[2] The importance of the indivisibility of plant and machinery may extend beyond the processing of the main product of a firm. In some industries, things which are waste materials from the point of view of the main product are capable of being made into valuable by-

Indivisibilities in Purchasing and Marketing. If a large firm can purchase a given quantity of some raw material for a specified total cost it does not necessarily follow that a small firm can purchase one tenth of this quantity of the material at one tenth of that total cost, or at the same cost per unit. Large-scale orders are important to the firms which receive them, discounts are frequently given, orders receive prompt consideration, and complaints are handled with dispatch and consideration. Indivisibilities in purchasing, and hence the economy of large-scale purchasing, may also be important in connection with the buying of power, transportation, fuel, and machines. Rates per unit for electricity, gas, and water may decline considerably as the quantity purchased increases, and freight rates for carload shipments are significantly lower than those for less-than-carload shipments.

In selling the final products of a firm, it is quite likely that the sales force required will not increase in proportion to production, especially if the increasing size of the firm enables it to deal more effectively with large-scale customers. Moreover, salesmen may be able to sell a family of related products, instead of a single product, without any great increase in total selling cost, and the related products may even help to sell each other. Finally, while some advertising can be used to cover only a local or limited market, other forms of advertising (such as, for example, advertising in magazines with a national circulation) cover a wide market and can be used most effectively only if the firm is large enough to be able to sell in such a market.

Indivisibilities in Finance and Research. Some financial advantages may be associated with the increasing size of the firm or productive unit. A large firm may be able to borrow funds at somewhat lower rates because it can offer larger and more readily marketable issues of securities, because its financial standing is more easily ascertained, and because the purchasers of securities feel that investments in a large firm are safer than those in a small one. It is also possible that a large firm may achieve greater stability of earning power than a small one, and it may be able to accumulate relatively greater financial reserves for business depressions or other emergencies. In connection with research, the indivisibilities are obvious. The cost involved in discovering a new product or a new process of production does not depend upon the scale on which it will be produced or used. Research activities which result in new processes, the elimination of pre-

products. Thus in the meat-packing industry, the horns, bones, blood, hair, and offal of slaughtered animals can be made into fertilizer, bone buttons, glue, brushes, pharmaceutical supplies, and many other products. The small-scale slaughterer, however, may not be able to make use of his waste materials, since the production of the by-products requires plant and equipment which cannot be set up on an extremely small scale, and since he does not have enough of the waste materials to permit him to make effective and economical use of the minimum practical size of this plant and equipment. The large-scale meat packer can use large and efficient sizes of the necessary plant and equipment. For such a producer the sale of by-products may make an important contribution to the earnings of the business.

viously necessary steps in manufacture, new products, or new uses for old products may be of great financial benefit to a firm, but a laboratory for research and investigation, furnished with the necessary scientific workers and equipment, costs a great deal of money. It may not be possible to set up such a research department on a very small scale, and the cost of the smallest practical size for such a department may be beyond the means of a small firm. A large firm may be able to make effective use of a research department, and it may be able to wait a long time, if necessary, for significant discoveries to be made and for the research department to begin to pay its way.

Indivisibilities in the Use of Labor. The learning power of the individual worker also seems to be imperfectly divisible. If he is asked to spread his learning power over a large number of tasks, the chances are that each task will be learned and performed rather ineffectively, while concentration on a single task will lead to thorough learning and efficient performance. For this reason, and for other reasons which we noted in Chapter V, the specialization of labor by tasks is able to make a contribution to increased efficiency and lowered cost, and a hundred workers with each worker performing some specialized task necessary to the fabrication of some final product will have a total output far more than ten times as great as that achieved by ten workers with each worker performing ten tasks. However, specialization by tasks cannot be fully developed unless the firm and its output are large enough to keep each worker busy at his specialized task; the result is that a large firm may be able to make more effective use of its labor than a small firm.

Turning now to labor of the managerial type, we find that a small firm may not be able to afford an extremely efficient and high-salaried manager because the firm is too small to permit the individual's unusual powers to be used fully and effectively, but a large firm may be able to employ the same individual economically. On the other hand, the efficiency of managerial employees of a small firm may suffer because each one has to perform too many types of managerial work, whereas in a large firm managerial work of the detailed and routine type can be divided among a number of specialized managerial workers.

Diseconomies of Extremely Large Size. Indivisibilities of the kinds just discussed are the factors responsible for the relatively low efficiency and high cost of small-scale firms in various industries. Conversely, it is the ability to make effective use of various indivisible productive agents or services that causes increased efficiency and lowered cost of production as a firm increases its size from small to what we have called optimum size. However, it is important to note that the economies of large size do not go on expanding indefinitely and without regard to the size to which a firm may grow. If they did, the tendency would be for production in each

industry to be concentrated in the hands of a single giant firm. In actual practice, as a firm goes on growing, some of the economies of large size will become stabilized and will increase no more, while others tend to turn into real diseconomies.

A few illustrations of this point may be given without covering again all the ground which we have just been over. A large quantity of a raw material may often be purchased at a lower cost per unit than a small quantity, but it does not follow that a still larger quantity can be purchased at an even lower cost per unit. Rates for gas, water, and electricity decline only up to a certain rate of consumption per individual or firm, and beyond that point decline no more. Railroad freight rates are lower for carload shipments than for less-than-carload lots, but are usually no lower for several carloads than for one, and so on. In all such cases, however, it should be noted that the prices or rates, having fallen with increased purchases, do not rise again as the volume of purchases is increased further. In other words, once a firm has become large enough to enjoy the lowest prices or rates, it suffers no disadvantage from becoming still larger.

A large machine may be more efficient in relation to costs of construction and operation than a small one, but it is not necessarily true that a gigantic machine will have a similar advantage over a large one. If an extremely large machine breaks down and needs repair, a considerable amount of productive capacity is tied up, while a similar breakdown on the part of one of several moderately large machines may not be nearly so serious. If machines become too large, it may be difficult for a firm to provide effectively for necessary fluctuations in output, and in some cases large machines may actually become cumbersome and difficult to operate.[3] However, it should be noted that, having developed its machines to their most effective size, a firm does not have to encounter any diseconomy in this respect as it grows larger. It can always use more machines of the most effective size instead of using larger machines.

Efficiency in the utilization of labor will also stop increasing at some point as a firm increases in size. The specialization of labor by tasks has been seen to be conducive to increased efficiency and lowered cost, but such specialization obviously cannot be carried beyond the point at which each worker has a single minimum task to perform. Once a firm has become large enough to take full advantage of specialization by tasks, it can make

[3] For example: "There is a story of a man who thought of getting the economy of large-scale production in plowing, and built a plow three times as long, three times as wide, and three times as deep as the ordinary plow and harnessed six horses to pull it, instead of two. To his surprise, the plow refused to budge, and to his greater surprise it finally took fifty horses to move the refractory machine. In this case, the resistance, which is the thing he did not want, increased faster than the surface area of the earth plowed, which was the thing he did want. Furthermore, when he increased his power to overcome this resistance, he multiplied the number of his power units instead of their size, which eliminated all chance of saving there, and since his units were horses, the fifty could not pull together as well as two."—J. M. Clark, *The Economics of Overhead Costs.* Chicago: University of Chicago Press, 1923, p. 116.

no further gains in this respect by increasing its size. On the other hand, further increases in size need involve no diseconomies in the utilization of ordinary labor, for several men can be set to work at each specialized task in a large concern.

Our discussion has suggested some of the reasons why we cannot expect a firm's efficiency to go on increasing and its average cost of production to go on falling as the firm expands in size beyond a certain point, but why should a continued expansion in size involve any actual diseconomies, losses of efficiency, and increases in cost? The answer to this question is found in the difficulties of management which arise as a firm becomes extremely large. To be sure, the detailed and routine work of management can be delegated successfully to a number of specialized managerial employees, but the same thing is not true of the most important type of managerial work—the making of ultimate decisions concerning the business of the firm. This highest type of managerial work, when placed in very few hands, becomes increasingly difficult to perform well as a firm increases in size beyond a certain point. Thus the very large firm is faced with a difficult problem of management. "At one extreme all authority may be delegated. Then there will be no unity of policy or uniformity of performance. At the other extreme, all decisions may be made by a final center. This system involves bureaucracy in its worst form: 'red tape,' hopeless delay, decisions based on diluted memoranda. Between these two extremes the large firm attempts to steer a middle course, but it never achieves that compactness, flexibility, and singleness of purpose which are possessed by every well-managed, medium-sized firm. The growing difficulty of coordination and decision-making eventually stops the growth of every firm."[4]

The Size of the Firm and Cost of Production. On the basis of the foregoing analysis, the relationship between the increasing size of the firm (accomplished by increasing all the factors of production) and the average cost of production of the product can be illustrated by means of a diagram, such as Figure 4. In this diagram, we have three "short-run" curves of average cost of the type which we presented in Chapter VI. The line AC^1 is the average cost curve for a small-sized firm in a particular industry (that is, a firm which has only a small quantity of those productive agents which are regarded as fixed at any given time) as different quantities of the variable agents are combined with the fixed agents. The lines AC^2 and AC^3 are average cost curves for successively larger sizes of the same firm.

As the firm increases in size from that represented by AC^1 to that represented by AC^2, the average cost of production of its product at normal output (lowest point of average cost) declines. This result occurs because the larger firm is able to make more effective use of the various imperfectly

[4] G. J. Stigler, *The Theory of Competitive Price*. New York: The Macmillan Company, 1942, p. 138.

divisible productive factors or services which we have discussed. However, as the firm increases in size from that represented by AC^2 to that represented by AC^3, average cost of production at normal output increases. That is, with this increase in the size of the firm, the advantages of large size remain stable or increase only slightly, while the grave difficulties of ultimate management produce a tendency to diseconomy and increasing

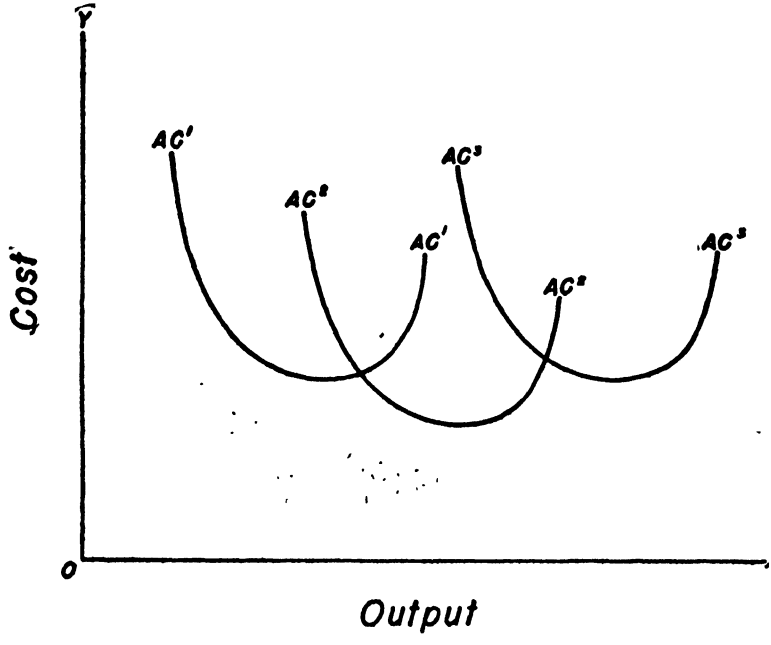

FIGURE 4.—Average Cost Curves in Relation to Size of Firm

average cost. This knowledge of the relationship between average cost of production and the size of the firm will be of assistance later on when we attempt to draw a "long-run" average cost curve for the individual firm.

Optimum Firms and Types of Industries. Our discussion indicates the error of the common notion which holds that increases in the size of the firm bring advantages with respect to efficiency and cost in some industries but not in others. In any industry, such advantages are obtained as the firm expands from small size to optimum size, while further increases in size bring diseconomies. However, while the optimum firm in any industry has greater efficiency and lower average cost than firms of any other size, industries do differ significantly with respect to the size of their optimum firms. For example, we should expect the optimum firm to be rather large in industries producing such complicated products as typewriters, adding machines, and automobiles, or articles such as sugar, salt, and meat, which

are standardized or graded articles of widespread consumption. On the other hand, if an article has a rather small local market and a high cost of transportation (as in the case of bricks and artificial stone products) or requires the use of materials which are widely scattered geographically and are difficult to concentrate in one place because of transportation cost or perishability (as in the case of cider or cheese), the optimum productive unit may be expected to be rather small.[5]

Optimum Size of the Firm through Time. Our discussion of optimum size of the firm should not leave the impression that all the firms in any industry will be exactly at the optimum size at any given time. In any industry it is quite likely that some firms will be larger than the optimum at any given time, while others will be smaller. However, if an industry is competitive, firms which are very much larger or smaller than the optimum may be eliminated sooner or later because of their low efficiency and high cost, while others which more closely approximate the optimum size will be able to remain in business but will operate less successfully than those of optimum size. As a result, over a long period of time and in a given state of the arts of production, all the firms in a competitive industry would be expected to move toward the optimum size, although this tendency would be incompletely worked out at any particular point in time. In non-competitive industries the forces which tend to drive the individual firms to approximate the optimum size would be much less strong than in competitive industries.

In any case, however, before all the firms in a given industry could adjust themselves to the optimum size, technological developments and changes in methods of production would be likely to bring about a change in the optimum size itself. Another factor which may have an important influence on the question of optimum size is the general level of production and economic activity in the economic system as a whole. Thus a recent study concluded that, at the bottom of the great depression in 1932, only the largest corporations (those having assets above 50 million dollars) were able to retain a slender margin of profit.[6] By 1941, when production and economic activity in general had ascended to higher levels in the United States, all sizes of corporations made profits, and the highest rates of profit were made by firms with assets of just under a million dollars.[7]

While it is very difficult to determine exactly what the optimum size of the firm is in any industry over a period of time, it is well known that the average size of the firm has changed through time in most American

[5] W. L. Thorp, *The Integration of Industrial Operation*. Washington, D. C., Department of Commerce, Bureau of the Census, 1924, pp. 88–89.

[6] Stockholders' equity in the corporations was used as the basis for expressing earnings as a rate of profit, and the earnings were computed before the payment of taxes.

[7] J. L. McConnell, "Corporate Earnings by Size of Firm," *Survey of Current Business*, May, 1945, pp. 6–7.

industries. For example, establishments with an annual product valued at $5,000 or more in all manufacturing industries in the United States showed an increase in the average number of wage earners per establishment from 37.3 to 51.4 in the years from 1914 to 1937. The average value of products per establishment increased from $132,823 to $287,279 over the same period, and the index of the volume of production per establishment increased from 100 to 185.[8]

Not only has the average size of the firm in manufacturing as a whole been increasing but the number and importance of large-scale firms have also been growing through time. Firms with an annual product valued at a million dollars or more included only 2.2 per cent of the total number of firms in manufacturing in the United States in 1914, but these few large firms employed 35.3 per cent of all workers in manufacturing and turned out 48.7 per cent of all manufactured products by value. In 1939, a quarter of a century later, the firms with an annual product valued at a million dollars or more included 5.2 per cent of all firms in manufacturing, employed 55 per cent of all manufacturing workers, and produced 67.5 per cent of all manufactured products by value.[9] Again, firms employing 1,000 or more workers increased only from .12 to .17 per cent of the total number of firms in all American industries from 1939 to 1943, but the employees of these firms increased from 30.46 to 44.69 per cent of all employees and the wages paid by these firms increased from 35.77 to 52.30 per cent of total taxable wages.[10]

The Practical Determination of Optimum Size. Thus far we have been tacitly assuming that there is some one size in any industry at which a firm can make most effective use of the various imperfectly divisible productive agents or services. As a matter of fact, it is often anything but a simple matter to decide on the optimum size when all factors affecting efficiency and cost are taken into consideration. For example, in a certain industry, a plant with an annual capacity of 100,000 units of product might be best from the point of view of technical efficiency in the operation of the plant, while a plant twice as large might be best from the point of view of efficiency of management. Similarly, a plant with a capacity of 500,000 units might be best for the purpose of marketing the product, while one capable

[8] Temporary National Economic Committee, Monograph No. 27, *The Structure of Industry*. Washington, D. C.: Government Printing Office, 1941, p. 5. In evaluating these statistics, we should remember that establishments with an annual product valued at less than $5,000 were not covered by the data, that increases in the factors mentioned were not at all regular and uniform over the period, and that data for *all* manufacturing industries together cover up the fact that in some individual industries the average size of the establishment remained constant or declined somewhat over the period, while in others it increased very rapidly. Thus in a number of industries the average size of the establishment declined by over 50 per cent from 1914 to 1937, while in others it increased by over 200 per cent.

[9] *Statistical Abstract of the United States*, 1937, p. 739, and 1939, p. 846.

[10] D. W. Paden, "Industrial Concentration of Employment," *Survey of Current Business*, April, 1945, p. 13.

of producing 1,000,000 units annually might be most desirable in order to facilitate financing through the sale of large issues of securities, or to lessen risk and increase stability during the ups and downs of business fluctuations.

Two courses of action may be open to the firm in such a situation. First, it may try to expand a single productive establishment to the size which will afford the most advantageous compromise among these various optimum sizes. This is frequently attempted, but it may be very difficult to hit upon the size at which the various optima will be most completely reconciled, and almost any size selected may be relatively unsatisfactory if the various optimum sizes are widely separated. Second, the firm may attempt to reach a solution by forming a combination of productive establishments. For example, it may be found that, from the point of view of technical efficiency in operation, although no further advantages may be secured as a plant expands beyond a capacity of 100,000 units of product per year, there will be no loss of the advantages already secured if such a further expansion takes place. Thus a combination of establishments could be formed which would consist, say, of three or four plants, each with a capacity of 200,000 units of product per year, so as to secure maximum efficiency of firm management while retaining maximum technical efficiency within each plant. All the plants would then engage in marketing and financing activities as a single concern so as to approximate the optimum sizes for these purposes. Such simple combinations of like productive units are very common in our economic system, although, as we shall see, there are other varieties of business combinations, and not all combinations are formed for the purpose of securing optimum size for efficiency in production.

Size of the Firm in Relation to Economic Systems. Before turning to a detailed study of business combinations, we should relate the subject which we have been discussing to differences in economic systems. Since the tendency to increased efficiency and lowered cost as a firm expands from small size to optimum size depends upon the imperfect divisibility of various productive agents or services, and since the tendency to decreased efficiency and higher cost as a firm expands beyond the optimum size depends upon the difficulties of decision-making and coordination, we must expect these tendencies to prevail in any economic system. The indivisibility of productive agents does not depend upon whether they are owned by the government or by private enterprisers, and the difficulties of large-scale management do not depend upon whether the managing is done in the name of the government or in that of private enterprisers. To be sure, in a socialistic economy in which productive establishments are owned and operated by the government, there may be no question of raising funds by means of security issues, government-owned plants may not be exposed to any danger of failure in connection with business fluctuations, and there

may be no problem of finding the size at which a productive unit can best take advantage of national advertising. By and large, however, the economies found in expanding the productive unit to optimum size do not depend upon the existence of any particular set of economic institutions, and we should expect them to exist under socialism, communism, or fascism about as well as under capitalism. And the same thing is true of the diseconomies of overexpansion, although some economic systems, in the absence of competition, might find it difficult to locate the optimum size.

COMBINATIONS OF PRODUCTIVE UNITS

The Number and Importance of Business Combinations. Combinations of productive units play a rather important part in American economic life. In 1919, for example, it was found that 7.8 per cent of all establishments in manufacturing and mining were members of combinations. There were 5,838 combinations in these fields and they controlled 21,464 establishments.[11] In 1937 there were 5,625 combinations in manufacturing alone. They controlled 25,699 out of 166,794 manufacturing establishments, or 15.4 per cent of the total number. Furthermore, they employed 51.1 per cent of the total number of wage earners in manufacturing, paid 55.3 per cent of the total wage bill, and produced 61.1 per cent of all manufactured products by value.[12] And these data give only an incomplete picture of the importance of business combinations, for combinations are of considerable importance in other fields of economic activity—as, for example, in mining and other extractive activities, in distributive operations (marketing of goods), and in the furnishing of various services (operation of hotels, dry-cleaning establishments, shoe-repair shops).

A combination is said to exist when two or more plants are operated or controlled by one ownership interest. A combination does not necessarily involve the putting together of productive establishments which formerly operated independently. Many combinations have come into being in this way, but a combination is also said to exist if a firm, which formerly operated only a single plant, constructs one or more additional plants in order to enlarge its business. Several methods are available for combining formerly independent plants in order to create a business combination. In some cases, holding companies are set up which exist for the purpose of owning and controlling the securities of other corporations. With its own securities or with cash, such a holding company buys up at least a controlling interest in the voting stock of the companies which are to be combined, and thus is able to control management, output, and prices. In other cases, actual mergers take place. While the holding company controls

[11] W. L. Thorp, *The Integration of Industrial Operations*, p. 107.

[12] Temporary National Economic Committee, Monograph No. 27, *The Structure of Industry*, p. 111.

two or more productive establishments through security ownership, the merger exercises the same control through holding the physical properties of the firms. The merging companies may lose their identity in a new company organized to manage all the properties of the old firms, or one of the old companies may remain in existence while the others are merged in it.

Finally, under our definition, a combination would be said to exist if a firm, with one plant of its own, leased and operated one or more other plants. However, we do not include as combinations any groups of firms which are controlled by means of such informal and intangible devices as communities of interest, gentlemen's agreements, common banking connections, price leadership, and so on. We should also note at this point that there are no geographical limitations in deciding when a combination exists. If a firm with a single plant in Chicago buys out, constructs, or otherwise comes to control another plant in New Orleans or New York, the result is a combination, but the same thing would be true if the additional plant were in Illinois, or Chicago, or even directly across the street from the first plant.

Simple Horizontal Combinations. The most common type of business combination is the simple horizontal combination, which exists when the two or more plants operated or controlled by one ownership interest turn out exactly the same product or products. Examples would be found in the operation, under one central management or control, of two or more like shoe factories, distilleries, coal mines, automobile factories, or paper mills; or in companies which operate or control a chain of hotels, grocery stores, bakeries, restaurants, or department stores. In 1937, over 63 per cent of all the combinations in manufacturing were of this type, and the extent of these simple horizontal combinations increased to 85 per cent of the total if one counted in all combinations which, while complex in total structure, had some simple horizontal elements (as in the case of a combination which controls ten establishments in the ice-cream industry and one in the cheese industry).[13] Simple horizontal combinations in manufacturing are found most frequently in such fields as food and kindred products; stone, clay, and glass products; products of petroleum and coal; and printing, publishing, and allied industries.

Simple horizontal combinations are sometimes formed, as we have seen, in order that the firm may approach the optimum size from several points of view, but there may also be certain interplant advantages in a combination which could not be secured by a single large plant. Any trade secrets, machines, or processes which are used in any one of the plants may be enjoyed by all the plants in a combination. This, of course, would not be true if the plants operated independently. The plants in a combination may

[13] *Ibid.*, p. 163.

be able to avoid a great deal of competitive duplication of plant and equipment, eliminate a number of formerly competing brands of goods, and make substantial savings in connection with advertising and selling expenses. Amid the fluctuations which occur in business conditions a combination may achieve greater stability than a single plant. If the plants in a combination are scattered geographically, they may not all suffer at the same time or to the same extent from adverse business conditions. Even if the plants are concentrated in one area, it may be possible to close down one or more of them in a period of depression and keep the others running at full capacity and hence economically, while, if the plants were independent, they might all find it necessary to run at part capacity and hence uneconomically. Finally, combinations with scattered plants will usually be able to fill orders from the nearest plant and thus save much of the "cross-hauling" of goods which would occur if the various plants were independently operated.

Vertical Combinations. In 1937, about 10 per cent of the business combinations in manufacturing were of the so-called vertical type;[14] that is, the two or more plants operated or controlled by one ownership interest functioned at different stages of the process necessary to prepare a final product for the market. For example, "the larger steel companies own and operate iron ore mines, quarry the necessary limestone, mine their own coal and convert it into by-product coke for use in their blast-furnace operations, roll their own steel and convert it into more advanced manufactures, and, finally, maintain a selling organization for the distribution of their products."[15] Such a combination, of course, extends beyond the field of manufacturing proper. Vertical combinations in the textile industry may produce, in separate plants, yarns, gray goods, finished cloth, clothing, and other fabricated textile products, while those in the paper industry may turn out lumber and timber products, pulp, paper, and highly finished paper products. In order to be considered vertical a combination need not handle all or most of the stages of a productive process in its own establishments. Any two stages are enough. Thus a manufacturing plant which operates its own retail stores, or produces its own raw material, is called a vertical combination if the two stages are carried on in separate plants. The four fields of manufacturing having the largest number of vertical combinations are forest products, textiles, paper and allied products, and iron and steel.

Several reasons may exist for the formation of vertical combinations.

1) At times, when suppliers of materials or distributors of final products are making large profits, it may seem desirable for a manufacturing concern to take over these operations. However, from the long-run point of view, it is doubtful just how much can be gained in this way. "As the

[14] *Ibid.*, p. 196.
[15] *Ibid.*, p. 192

integrated concern must have invested capital in all the stages of production, there must, in the long period, be a return to that capital at least equal to the return which could be obtained by investment elsewhere." [16]

2) A vertical combination may be better able than several independent concerns to coordinate the various stages or steps in the productive process. It may be possible to plan production so that the output of each plant will exactly meet the requirements of the establishment at the next stage of operations. Inventories of various types of goods may be reduced and the need for working capital lessened.

3) A vertical combination may achieve stability and independence of operation. The establishments at all stages of such a combination may be assured of an adequate and steady supply of exactly the right type and quality of materials or other goods from the establishments at the next lower stage in the productive process, and a uniform and specialized labor force may be maintained. By carrying on operations at various levels or stages of the productive process, a more favorable distribution of risk, a greater degree of stability with respect to business fluctuations, and increased borrowing power may be secured by a vertical combination in comparison with a single plant producing only the final product.

4) "Economies arising from technical advantages may be realized through continuous plant operations. Thus, in the combination of blast furnaces and steel mills there may be considerable savings in fuel by avoiding the necessity of reheating the metal; while in a combination of pulp and paper mills it is possible to transfer the pulp direct to the paper mill without the expense of drying." [17]

5) Finally, it may be possible for a vertical combination to derive economies by spreading administrative or managerial costs over a large organization, and considerable savings may be realized in connection with selling operations, advertising, and bad debts.

Complex Horizontal Combinations. Other combinations of productive units are often called complex horizontal combinations. In such cases, the two or more plants operated by one ownership interest turn out unlike but not successive products. Although the various plants produce different articles, the plants cannot be said to be located at different levels or stages of the productive process necessary to the completion of some final product. A complex horizontal combination may produce several products from the same material (bowling alleys and pins, billiard tables, and phonographs); main products and by-products, as in the case of combinations which engage in slaughtering and meat-packing as the main line of work but produce in separate plants such by-products as fertilizer, glue and gelatin, soap, shortening, and sausage casings; or even unrelated products,

[16] *Ibid.*, p. 194.
[17] *Ibid.*, p. 195.

as in the case of a company in the explosives and ammunition business which controls plants producing condensed and evaporated milk, pulp, tallow and grease, and tanning materials. It is sometimes very difficult to account for the existence of complex horizontal combinations producing unrelated products.

Combinations of Combinations. The foregoing classification of business combinations may seem reasonably clear cut, but in practice there is much overlapping among the various types of combinations, and there is frequently much doubt as to the exact group in which a given combination should be classed. Moreover, some complex combinations actually fall in all three major groups. That is, a concern such as General Motors is a vertical combination operating at many successive levels or stages of the productive process necessary to the turning out of automobiles, but at various levels it operates several plants turning out exactly the same product or performing the same function, and it produces several products (home units for producing electricity, for example) which are not very closely related to its main product.

The Success of the Combinations. In actual practice, business combinations in the United States have not been nearly so efficient and profitable as the discussion of their possible advantages would lead one to expect. One study of thirty-five combinations concluded that the earnings of most of the combinations in their first decade of operation failed to reach the expectations of their organizers or promoters and, for the thirty-five combinations as a whole, were nearly one fifth less than the earnings which the separate productive units had enjoyed before they were combined.[18] This result occurred in spite of the fact that some of the combinations had made large additions to plant and equipment by means of new financing. Another investigation, which covered from twenty-nine to forty-eight combinations in the various years of the period 1900–1913, decided that none of the combinations made extraordinary earnings in any of the years and that few of them had even reached earnings of 10 per cent, while many had returned 2.5 per cent or less on their nominal net worth. In the case of twenty-three out of forty-four combinations, the common stocks were higher in price in 1914 than in 1900, while the reverse was true in the case of the other twenty-one combinations.[19]

Several factors are sometimes advanced in explanation of the relative lack of success of many business combinations. In the first place, many combinations apparently were not formed for the purpose of promoting increased efficiency in production. Some were created in the hope of suppressing competition and achieving monopoly control over a certain product

[18] A. S. Dewing, "A Statistical Test of the Success of Consolidations," *Quarterly Journal of Economics*, November, 1921, pp. 84–101.

[19] *Mergers in Industry*. New York: The National Industrial Conference Board, 1929, p. 39.

or group of products, although this hope was not always realized in practice. In other cases the desires of promoters and investment bankers to make profits by organizing the combinations and selling their securities seem to have been dominant. If combinations were to be formed to realize certain economies and increase efficiency, their organization would be expected to occur in times of business depression when these advantages would be most needed. Instead, most of the combinations have been formed in prosperous periods when their securities could be readily marketed, and the combinations have often been reckless about including inefficient, high-cost concerns. If combinations have not been formed to increase efficiency, it is hardly to be expected that they will attain great efficiency by accident.

In the second place, it may be said that intraplant economies are more important than interplant economies from the point of view of increasing efficiency and lowering cost. There is usually little that combinations can do to increase intraplant economies, and combinations may grow to such great size in seeking interplant economies (or in trying to attain control of the market) that opportunities for economies within the individual plants are neglected or are actually lost. Finally, the most important fact seems to be that combinations are often unwieldy from the point of view of attaining efficiency in management. We are referring here, of course, to the problem of managing a combination as a whole as well as to that of management within the individual plants.

The problems of management are severe even when a combination is of the simple horizontal variety, but they are much more difficult when a combination is complex and operates many types of businesses at the same time. Consider, for example, the combination which operates "businesses producing numerous breakfast cereals, animal feeds, gelatin, ice-cream mixtures, a medley of desserts, a coffee and tea business, a cake and bread business, a chocolate and cocoa business, a coconut-meat business, a syrup business, a nut business, a salt business, a baking-powder business, a business of manufacturing laundry aids, an oyster business, a business of producing frosted foods, a business of processing corn products, a business of canning fruits and vegetables, a business of manufacturing cottons and shipping cases and bags, a business of manufacturing tin cans and a business of meat packing." [20] How could any man or small group of men be familiar with all of these businesses and be able to administer and coordinate them effectively from the headquarters of the combination?

The ultimate management of large and complex combinations must necessarily be absentee management, and the members of their boards of directors frequently do not confine themselves to the affairs of any one combination. Instead, they are members of the boards of directors of sev-

[20] Temporary National Economic Committee, Monograph No. 13, *Relative Efficiency of Large, Medium-Sized, and Small Business*, p. 119.

eral combinations or companies and cannot possibly be fully informed about the affairs of all the concerns which they are supposed to help manage. As a result, they often confine themselves in practice to the financing activities of the various companies and neglect the more fundamental aspects of company management. This situation is not conducive to the attainment of maximum efficiency in the individual productive units or plants. On the whole, then, the conclusion is that many combinations have been carried well beyond the point at which most effective use could be made of the various indivisible productive agents or services, and almost insuperable difficulties of management have been the result.

MONOPOLIES

Concentration in Industry. In some industries in the United States the process of concentration has gone far beyond the setting up of single large scale productive units or even combinations of productive units. For example, there were said to be nine fields of production in the United States in which one firm controlled all or practically all of the nation's supply of an individual product in 1937. In nine other cases, pairs of firms controlled all or nearly all of the supply of certain economic goods. In the case of 37 products, four firms accounted for the entire supply. In 164 cases, four firms produced over 90 per cent of the supply; and in 328 other cases the part played by the four leading firms was not disclosed (in order to prevent the identification of individual firms). The general conclusion was that from two fifths to one half of all the goods under consideration were produced in fields in which four concerns controlled 75 per cent or more of the supply.[21]

The Development of Monopoly. There is no doubt of the existence of effective monopoly in any industry in which one firm controls all or practically all of the production of a given economic good, but an effective degree of monopoly power may also exist in an industry in which a few large firms control all or most of the productive facilities. That is, such a small number of large firms, *by acting together and regulating production or other activities,* may be able to control the prices charged throughout the industry for their product or products. In such cases it is not necessary for the few large firms to enter into a formal combination in order to exercise monopoly power in an industry. Sometimes the firms, while remaining unassociated in other respects, will enter into actual agreements to establish and maintain uniform prices and terms of sale. In other cases, the more informal device of price leadership is used, with one firm acting as the leader in establishing or changing prices while the others faithfully follow suit.

[21] Temporary National Economic Committee, Monograph No. 21, *Competition and Monopoly in American Industry.* Washington, Government Printing Office, 1940, pp. 69, 98.

In some industries a system of "delivered prices" is maintained through the use of one or more "basing points." When several basing points are used, all firms in a given district charge their customers a uniform base price, plus freight from the basing point to destination. Frequently the actual freight from seller to customer is less than that from the basing point to the customer, and the difference goes into the seller's profit. When making sales in that district, plants located outside a given basing-point district must ordinarily charge the same prices as those inside the district, and absorb the higher freight charges as best they can. Thus a customer in a given district would be quoted the same price by all firms in the district, and by those outside the district as well, so that price competition is eliminated. The firms which operate basing-point systems often claim that these uniform prices are only "asking prices," and that actual prices charged and received may differ somewhat from one firm to another. However, basing-point systems are usually regarded as a device for maintaining monopoly control.

Effective monopoly has been maintained in some industries by means of patent pools. When important patents are owned by a small number of large firms, each firm may grant licenses to the others to use its patents, or all firms may pool their patents. This group of firms may then use its resources to exclude new producers from the general field of production, by refusing to grant licenses to outsiders or by charging very high royalties for the use of patents. When licenses are granted to new firms, the members of the pool may attempt to control output, markets, or prices charged by such newcomers.

In addition to the methods of maintaining effective monopoly which have already been described, market sharing is sometimes practiced by a few large firms in a given industry. Market sharing means simply that the firms do not compete against each other for the same customers. Each firm has a particular share of the general market and works it exclusively. Again, control in some industries has been achieved through interlocking directorates and through the relationships of the firms with a common financial organization. Finally, the activities of trade associations or industrial institutes in some industries have operated to promote or maintain effective monopoly.

Monopoly Prices. The social and economic significance of monopolies, and the efforts of government to deal with the monopoly problem in the United States, would furnish ample material for a complete book. In the present chapter we can only indicate some of the bases for the popular attitude of suspicion and distrust toward monopolies which has long existed in this country.

In general, it is not necessary for a business organization to reach the monopoly level in order to enjoy to the full the advantages of large-scale

production and business combination, and there is little to indicate that monopolies make large earnings because they have low costs. Rather, their success has almost always been due to the restriction of output and the charging of prices which were well above the level of cost of production. In this way the monopolies have made large profits, while the consumers have been given smaller quantities of economic goods to enjoy than they might have had under competitive conditions, and their pocketbooks have been hard hit by the high prices charged by monopolies.

It is, of course, very difficult to determine the *exact* effect of monopoly control over prices. To do this, we should have to compare the monopoly price with the price which would have prevailed under competitive conditions. Since there is no way to determine the latter, this comparison cannot be made. We know, however, that the prices charged under monopoly have usually been sufficiently high to yield large profits. It is true, as some defenders of monopolies have said, that prices in certain monopolistic industries have been stable or have even declined at times, but it is also true that monopoly profits may be gained without price increases. A stable price will yield large profits if it is high enough in the first place or if it is accompanied, over a period of time, by falling costs of production. Even falling prices over a period of time will be highly profitable if costs fall more rapidly than prices.

The profits made by monopolies are a fair general indication of the effect of monopoly control upon prices. For example, among the early trusts, the original Standard Oil Company had earnings which ranged between 48.8 and 84.5 per cent on its investment, and averaged 61 per cent between 1896 and 1906.[22] Again, the Aluminum Company of America made a net income of $335,000,000 over a fifty-year period, though its original investment was only $2,000,000.[23] It seems clear that such profits can be made only from prices which are well above the cost of production level, whatever may happen to the absolute level of the prices. From the point of view of society, then, one of the principal objections to monopoly has been its effect on prices and the burden thus placed on the consumers of monopoly products.

Unfair Tactics of Monopolies. When the smaller firms in an industry dominated by a monopolistic organization have refused to cooperate, or when new competing firms have sprung up, the monopolistic firms have often resorted to unfair tactics to attain their ends. Those monopolistic firms with large or national markets have often engaged in local price cutting. That is, in areas in which they have had competition they have cut prices to a level below cost in order to force their competitors to take

[22] *Ibid.*, p. 66.
[23] *Ibid.*, p. 72.

losses. Meanwhile, they themselves have been able to break even or make money by continuing to charge high prices in noncompetitive areas.

Firms that controlled a tremendous volume of business were once able, by playing one railroad against another, to secure rebates on freight rates, or to induce the railroads to establish particularly favorable rates out of the cities from which these firms shipped most of their products. Indeed, a large company once made an arrangement whereby it not only received a substantial rebate on its own shipments but also received a similar sum from the freight payments made by its competitors. The unfairness of such a policy is apparent.

Monopolistic firms have sometimes ordered dealers selling their products to refrain from handling the products of competing companies, thus limiting the market open to competitors. A similar device was the "tying contract." A large machinery company, for example, leased indispensable machines on which it held patents, but required the lessees to contract also for the use of other machines on which the patents had expired, and which were offered by other firms at lower figures. In this way, the company forced manufacturers to use its machinery exclusively, and thus destroyed competition.

Sometimes monopolistic firms have interfered with the flow of services and supplies needed by their competitors. They have persuaded or forced banks to cut off credit, and to call in the loans of competing firms, and have led newspapers and periodicals to refuse competitive advertising. Railroads have been induced to develop sudden "shortages" of freight cars of the necessary type, and sellers of raw materials have been persuaded not to fill the orders of competitors, or to fill them with inferior materials or only after long delay.

When public feeling against certain monopolies has run high, they have often made it appear that firms under their control were independent. Those customers who might refuse to buy from the monopoly would buy from these "bogus independents," and these companies could compete with the genuine independents without their fraudulent nature becoming known. Monopolistic firms have brought spurious lawsuits against competitors, charging patent infringement or other injury. These suits would often tie up the business of the competitors and involve them in heavy legal costs, even though the suits were eventually dismissed. Competitors have sometimes been intimidated by the mere threat of such suits. On the other hand, monopolistic firms have infringed upon the patent rights of competitors, or appropriated trade secrets gained through bribery or espionage, confident that they would make more through these illegitimate acts than any amounts which the courts might award the competitors by way of damages. Customers of competitors have been bribed to cancel

orders or to default on payments, and valued employees have been bribed to leave the employ of competitors.

Finally, monopolistic concerns have sought to damage their competitors' products or reputation. For instance, a large machinery manufacturer is said to have "doctored" machines made by competitors so that they would break down or fail to work satisfactorily. The same company bought up competitors' machines and offered them for resale, displaying them as "junk" or as "inferior products" at 30 cents on the dollar.

Monopolies and Business Cycles. Monopolies have also been charged with contributing to the instability of our economic system. In periods of good business they maintain stable prices for their products in the face of increasing production and favorable unit costs, and without increasing the wages of labor significantly. Such policies, in spite of all that can be done to support consumption through installment selling or other credit devices, are likely to lead to business breakdown and depression in the end. And in periods of depression, the monopolies are likely to hold their prices at about the customary level, taking losses by selling fewer units at the stable price rather than by selling more at a lower price. This policy results in laying off labor, curtailing purchases of materials and supplies, and reducing consumer purchasing power, and eventually affects all industries and adds to the cumulative downward spiral of depression. Some economists hold that cyclical swings in business can never be eliminated, or even greatly reduced in severity, so long as a considerable number of major industries are controlled by monopolies with their rigid or "sticky" prices.

Monopolies and Economic Progress. Monopolies may also operate as a bar to economic progress. Virtually complete monopolies need not worry about developing new methods and processes or making new inventions, for their profits do not depend upon continuous progress in production. Indeed, they may even go to the other extreme and acquire control of new inventions by fair means or foul, only to put them on the shelf for a long time so that they will not overturn existing conditions in the industry. However, in certain industries which are dominated by a few large concerns, there has often been keen competition in the field of technological improvement; and the public has frequently benefited by this sort of competition.

Local and Other Monopolies. The general distrust of monopolies in this country is by no means confined to the great industrial monopolies which operate on a nation-wide scale. It may frequently be true that consumers are more seriously affected by local monopolies, such as public utility companies (in spite of governmental authorization and regulation) or organizations of, say, coal dealers, than by the great industrial monopolies. Under certain conditions, organizations of workers may increase in power until they are virtually monopolies, and such organizations may also affect

consumers with regard to the prices which must be paid for the products of industry. Even the operations of agriculture under the auspices of the Agricultural Adjustment Administration have had some of the characteristics of monopoly.

The powers of monopolies in dealing with the public are not unlimited. If the price charged for the product of a monopolist is excessive, it is possible in some cases that someone will develop another product which can be substituted for that of the monopolist. Extremely high profits on the part of a monopoly may induce other business enterprises to invade the monopolized field and set up concerns which will be able to furnish real competition for the erstwhile monopolist. Indeed, even though the monopolist has the worst of intentions, he may lack the knowledge or ability to exploit his monopoly position to the fullest extent. However, these factors are often ineffective and in general the people of the United States have not been willing to rely on them entirely. Instead, repeated attempts have been made to control the formation and operation of monopolies through antitrust legislation, but on the whole the results have been far from completely satisfactory.

QUESTIONS AND PROBLEMS

1. "The tendency to increasing efficiency and decreasing average cost as a firm expands in size toward the optimum results from the fact that the individual agents of production are imperfectly divisible." Explain.

2. "The indivisibility of productive agents can be clearly seen in connection with such things as a firm's machinery and plant." Show whether you agree.

3. Comment on indivisibilities in purchasing and marketing.

4. In what sense is labor an indivisible agent of production? Explain.

5. Why may a firm of extremely large size stop securing additional economies or even run into important diseconomies? Explain.

6. "Large-scale concerns are more successful and profitable than small-scale concerns in any given field of production." Do you agree? Explain.

7. "The optimum size of the firm is quite constant from one industry to another." Show whether you agree.

8. "The optimum size of the firm is rather stable through time in any given industry." Discuss.

9. "At any given time it is quite likely that all the firms in a competitive industry will be of optimum size." Do you agree? Explain.

10. "The 'optimum size of a productive unit' is a term which may have a variety of meanings." Explain.

11. Distinguish between simple horizontal and vertical combinations.

12. "A simple horizontal combination is formed either when one hotel buys and operates another or when it constructs a second hotel for itself." Do you agree? Explain.

13. "The chief purpose in forming simple horizontal combinations is usually to secure the advantages of large-scale production to a greater extent than

would be possible for a single productive establishment." Show whether you agree.

14. "A shoe factory may form a vertical combination by acquiring and operating either a retail shoe store or a tannery." Explain.

15. "The advantages of vertical combinations are quite different from those of simple horizontal combinations." Do you agree? Explain.

16. "Combinations of business units are seldom more profitable than simple large-scale productive units." Discuss.

17. "Monopolies are usually successful because they are able to produce more efficiently than ordinary combinations or single large-scale productive units." Show whether you agree.

18. Why has the public attitude toward monopolies been almost invariably unfavorable?

19. Are there any checks on the powers of monopolies? Explain.

REFERENCES FOR FURTHER READING

Big Business—Its Growth and Its Place. New York: The Twentieth Century Fund, 1937, Chaps. 3, 8.

Fetter, F. A., *The Masquerade of Monopoly*. New York: Harcourt, Brace and Company, 1931, Chaps. 1–17.

How Profitable Is Big Business. New York: The Twentieth Century Fund, 1937, Chaps. 2, 13, 14.

Robinson, E. A. G., *The Structure of Competitive Industry*. New York: Harcourt, Brace and Company, 1932, Chaps. 2, 3, 5, 7, 9.

Stigler, G. J., *The Theory of Competitive Price*. New York: The Macmillan Company, 1942, pp. 128–138.

Temporary National Economic Committee, Monograph No. 21, *Competition and Monopoly in American Industry*. Washington, D.C.: Government Printing Office, 1940.

Temporary National Economic Committee, Monograph No. 27, *The Structure of Industry*. Washington, D.C.: Government Printing Office, 1941.

VIII

The Form of the Productive Unit

In the preceding chapter, we examined the matter of the *size* of the productive unit or firm in relation to efficiency and cost. Now we must turn to the description and analysis of the various *forms* in which productive enterprises may be organized. The discussion will include five types of productive organizations—single proprietorships, partnerships, corporations, cooperative enterprises, and public or governmental enterprises. On the whole, the question of the form of the productive unit is closely related to that of its size. In many fields of production, the optimum size of the firm is relatively large and, as we shall see, some forms of the productive unit are better suited for the attainment of large size than others.

THE SINGLE PROPRIETORSHIP

The single proprietorship, the oldest form of the productive unit, still outnumbers all other varieties in our economic system. The single proprietorship is a one-man business in the sense that one man owns the business and assumes ultimate responsibility for carrying it on. This does not mean that the owner of the business furnishes all the management and labor necessary to its operation. The single proprietorship may employ a considerable number of workers and do a large volume of business, and the proprietor may delegate much of the detailed work of management to subordinates. However, the proprietor, as the owner of the business, is responsible for final decisions and is the sole controller of business policies.

The single proprietorship has some advantages as a form of the business unit. In general, businesses of this type are flexible. The proprietor may reach decisions quickly and act upon them with dispatch, since he has no partners or board of directors to consult. This quickness of action may be of considerable importance in some fields of economic activity, as, for example, in the production or marketing of goods which are subject to changes in style or fashion. Single proprietorships may be entered into with less formality than most of the other types of business units, and they may at times be relatively free from burdensome regulations and

special taxes which are applied to corporations. The single proprietor is his own boss and has an independence which means much to some individuals; that is, there are probably many people who would prefer to derive a modest income from businesses of which they are sole proprietors than to receive a somewhat larger income as employees of someone else. Whatever net income the individual proprietor is able to secure belongs to him alone and does not have to be shared with anyone else. This fact should furnish the single proprietor with the maximum incentive to efficiency.

However, the other side of the picture is somewhat depressing. The single individual will often neither have nor be able to borrow capital funds in large enough amounts to permit him to establish an enterprise of the most efficient size in many industries. He may be able to finance a farm or a retail store which can operate efficiently, but a large-scale factory or mill may be entirely out of reach. Moreover, an individual might be unwilling to set up a large-scale single proprietorship even if he were financially able to do so. The single proprietor is completely liable for the debts of his business. From the legal point of view, the wealth and income of the single proprietor as an individual and the wealth and income of his business enterprise are inseparable. If the enterprise is unsuccessful, the personal wealth of the proprietor (with limited exceptions) and his income from other sources than his business venture may be taken from him to satisfy the enterprise's creditors, even though he had not considered such wealth or income as being in any way connected with the business. Similarly, business assets could be seized to satisfy personal debts of the proprietor. Finally, the proprietor may have difficulty in finding a purchaser for his enterprise if he wishes to retire from it, and the life of the business terminates legally with his own.

Since the individual is likely to be unable or unwilling to set up a large-scale productive unit in the form of an individual proprietorship, we find that this form of the productive unit is of little importance in manufacturing, mining, transportation, public utilities, and other fields in which the optimum size of the productive unit is rather large. More than three fourths of the individual proprietorships in this country today are in agriculture, and the others are concentrated to a great extent in retailing and in service fields (doctors, dentists, lawyers). These fields are, of course, those in which the optimum size of the productive unit is fairly small or in which small enterprises are not at a great disadvantage in comparison with large ones. Individual proprietorships, considered as a whole, have been declining steadily in importance through time, even though they have remained very numerous. That is, they have employed a decreasing proportion of all workers and have produced a declining proportion of all goods produced.

THE PARTNERSHIP

The partnership is another old and well-known form of the productive unit. It may be defined as *an association of two or more individuals for the operation of any industrial, commercial, or other business undertaking, or for the pursuit of a certain occupation or calling.* In most cases the partnership operates on the basis of a written contract or agreement between the partners, which may specify the financial contributions to be made to the business by the various partners, the duties to be performed by each in connection with the venture, and the extent to which each is to share in the earnings of the business. The partnership, then, replaces the individual proprietor with a group of proprietors whose relations to each other in connection with the business are determined in advance.

Advantages of the Partnership. Partnerships may enjoy several advantages in business as compared with single proprietorships. The several individuals in a partnership will usually be able to furnish a business enterprise with a greater quantity of capital funds than a single proprietor could provide. Moreover, other things being equal, a partnership should have a greater borrowing power than an individual proprietorship, for the resources of all the partners will serve as security behind any loans which are obtained. In a partnership there are often several partners to consider important business problems and to give advice on business policies, while the single proprietor must stand or fall on the basis of his own business ability. In the operation of a business several heads are likely to be better than one.

A partnership may also prosper because each of the various partners has some specialized talent or ability to contribute to the business. In the case of factory operation, for example, one partner may be an expert in securing efficiency in the internal operation of the plant, another may give excellent service in purchasing all sorts of things for the business and in marketing its products, while a third may be adept at hiring and handling employees. In a professional partnership in the field of medicine, the various partners may specialize in individual parts of the field, as in obstetrics, surgery, diseases of children, or ailments of the ear, nose, and throat. Partners in a law firm may specialize in similar fashion, or in quite a different way. One partner may be unusually able in the preparation of cases for trial though he is of little use in the actual give-and-take of the courtroom, while a second partner may present the reverse of these talents. A third partner may not have great ability in either of these respects, but he may have the social and business connections which will produce a large volume of business for the firm. Specialization of such kinds is obviously impossible for the individual proprietor. While the partnership has some advantages in comparison with the individual proprietorship, it has very few in comparison

with the corporation. About all that can be said for the partnership in this connection is that it may be free from certain regulations and special taxes imposed on corporations and that its business may retain a sort of personal character which may be important in retailing and service enterprises.

Disadvantages of the Partnership. Although the partnership ordinarily has an advantage over the individual proprietorship in the matter of raising capital funds, it is usually at a disadvantage in this matter in comparison with the corporation and is likely to be unable to raise the funds necessary for enterprises which are really large scale. This is, of course, only a general statement, for some partnerships, such as J. P. Morgan and Company (which only recently changed to the corporate form), apparently never suffered from an insufficiency of capital. A partnership venture ordinarily terminates with the death or withdrawal of any partner. Either eventuality does not prevent the other partners from engaging in the same kind of business in the future, but the existing partnership is broken up. When a partner has died a new partnership cannot be formed until the interests of the deceased partner's heirs have been cared for. A partner cannot withdraw from a partnership at will by selling his interest to some other person. Instead, the old partnership must be terminated so that his share may be withdrawn. Here, again, these are only general statements, for there are considerable variations from one state to another in the laws affecting partnerships.

Another significant disadvantage of the partnership is found in the fact that the partners have unlimited liability with respect to the debts of the business. It is possible for the personal wealth of the partners to be taken for the satisfaction of debts of the partnership, and, if the other partners are unable to pay, any one partner may be held for all the debts of the business and not merely for his share as indicated by the partnership agreement. Moreover, obligations undertaken in the name of the business by any one partner are binding on the others. Two partners may agree definitely that one of them is to do all the purchasing for the business, but the other partner can go out and buy a dozen new machines and the first partner is also liable for the debt thus incurred. This liability of all partners for the acts of any partner may seem severe, but firms dealing with a partnership could hardly be expected to know all of the details of the partnership agreement and must therefore be allowed to assume that any partner is the authorized agent of the partnership.

In some states, a bit of relief from the burden of unlimited liability is provided by permitting a partnership to include one or more "limited partners," whose responsibility for debts is limited to their actual investment in the business. Limited partners cannot participate in the work of managing the business, cannot do business in the name of the partnership, and

cannot have their names appear in the name of the firm. Not all partners in a business venture may be limited, and the public must be given due notice that some of the partners are limited if they are to obtain relief in this way. Finally, a partnership may be at a disadvantage in comparison with the individual proprietorship in that the partners may have great difficulty in agreeing on matters of business policy and courses of action, and may therefore be slow in reaching and carrying out decisions. If the partners are completely unable to agree, dissolution of the partnership is the only solution.

Significance of the Partnership. Because of its serious disadvantages, the partnership form of business organization is not used very widely in our economic system. In fact there are about forty individual proprietorships for each partnership in the United States. While some partnerships are found in a large number of different fields of economic activity, they are concentrated to a considerable extent in such fields as wholesale and retail marketing, the professions and service production in general, and finance. In general they are likely to be found in fields in which enterprises can operate without a very large capital investment and are not likely to pile up heavy debts.

THE CORPORATION

Nature and Importance of the Corporation. The disadvantages of the individual proprietorship and partnership forms of business organization weigh so heavily that it is doubtful whether large-scale production, with its important potential economies in many industries, would ever have developed to any considerable extent had it not been for the corporation. The large-scale enterprise usually requires a heavy investment of capital funds, which individual proprietorships and partnerships are frequently unable to furnish or acquire. Moreover, even if they could acquire adequate amounts of capital funds, the unlimited liability and limited life of these forms of organization, as well as the difficulties of transferring ownership, would often discourage their use for large-scale enterprises. The corporation is strong where these other forms of organization are weak, and the use of the corporate form of organization has increased steadily as large-scale production has developed.

A corporation may be defined as *a group or association of individuals (stockholders) which, on the basis of a grant of power, known as a charter, received from a state government, operates as a single person in carrying on a business enterprise.* As a legal person, the corporation has most of the rights and powers of ordinary individuals in carrying on its business. It can own wealth, sue and be sued in the courts, borrow and lend money, and enter into contracts. Its charter may or may not limit closely the types of economic activity in which the corporation may engage. Since corporations

pay fees for their charters and certain types of taxes to the states which grant these charters, some states have entered into active competition to get large numbers of corporations to seek charters from them. This competition has resulted in the lifting of restrictions from corporations and the extension of corporate powers as provided by the charters. In some cases corporations have been granted the right to engage in virtually any type of business almost anywhere. It is common for corporations to incorporate in states which grant the most favorable charters and then maintain only "dummy offices" there, while their real central offices are located in New York, Chicago, or other centers of population.

Since many states require only a very small number of stockholders in a corporation, it is possible that a corporation may have no more stockholders than a partnership has partners, but many corporations have hundreds or thousands, and some have hundreds of thousands, of stockholders. The individual partners in a business venture ordinarily take an active part in the management of the enterprise, but the individual stockholders in a corporation usually do not and have no power to act in the name of the corporation. The stockholders have a nominal right to control and manage the affairs of a corporation, but their powers in this respect are normally delegated to a smaller group of persons known as the board of directors, who in turn pass on certain powers to the other officials and managers of the business organization.

The corporation is by far the most important form of business organization to be considered in the present chapter. It has been estimated that in 1937 some 60 to 65 per cent of all the economic activity in the United States was carried on by corporations. However, as we see in Table 6, corporations are much more important in some lines of economic activity than in others. In agriculture, for example, corporations were responsible for only 7 per cent of the total income produced, and in three other fields of activity they accounted for only 30 to 36 per cent of income produced. At the other extreme, corporations in the fields of manufacturing, mining, communications, and electricity and gas turned out 92, 96, 100, and 100 per cent, respectively, of the income produced in these branches of activity.

The Accumulation of Capital Funds. Several factors account for the rapid development and present importance of the corporation as a form of business enterprise. In the first place, the corporation is better able than other types of private enterprise to raise large sums of money for purposes of capital investment. Since the corporation is authorized to sell securities, known as stocks and bonds, which have a relatively small price per unit, it can reach more readily than the other types of enterprise the funds which have been laid aside by small savers. A man with only a hundred dollars or so to invest could not hope to set up an individual proprietorship in most fields of production, nor could he buy his way into many partnerships with

Table 6: Relative Importance of Various Branches of Economic Activity and the Percentage of Total Income Produced by Corporations in Each Branch, 1937*

Type of Activity	Per Cent of National Income Produced	Per Cent Produced by Corporations in Each Branch
Manufacturing	24.0	92
Government, including work-relief wages	13.5	58
Trade	12.5	58
Services (professions, amusements, hotels)	11.9	30
Finance (banking, insurance, etc.)	9.3	84
Agriculture	8.9	7
Transportation	7.3	89
Mining	2.1	96
Contract construction	2.1	36
Electric light and power and manufactured gas	1.6	100
Communications	1.3	100
Miscellaneous	4.2	33
	98.7	60–65

* Source: Temporary National Economic Committee, Monograph No. 20, *Taxation, Recovery, and Defense*, p. 319. Washington, D. C.: Government Printing Office, 1940.

such a sum, but through the purchase of one or more shares of stock he could become a part owner of almost any corporation in the country. Thus a large corporation may have many thousands of stockholders, no one of whom owns as much as 1 per cent of its total stock. This ability of the corporation to raise large amounts of capital funds is very important in fields of production in which the optimum size of the productive unit is relatively large, for it permits the corporate firm to expand to the size at which it can take full advantage of the various indivisible productive agents and services.

Corporate Securities. In general, corporations issue three major classes of securities, although there are many minor variations possible within these general divisions. The *bonds* of a corporation, owned by bondholders, are in reality the promissory notes of the enterprise. The bondholders are creditors, and not part owners, of the corporation. They have lent sums of money to the enterprise, and the corporation undertakes to return these sums at some definite time in the future. In the meantime interest is paid on the bonds at a fixed rate. If the enterprise is to remain solvent, this interest must be paid, whether earned by the corporation in a certain year or not, and the corporation usually pledges all or some part of its tangible assets as security for the bondholders. The bondholders cannot share in the profits of the business, if any are made, but they are quite sure of getting their fixed rate of interest. Since the bondholders are not part owners of the

corporation, they have no voice in its control or management so long as the corporation fulfills its obligations to them, but in many cases it is provided that the bondholders shall acquire voting power and partial or complete control of the business if the corporation fails to pay interest due on the bonds. Moreover, if the corporate venture is a failure and has to be liquidated, after ordinary business debts of the enterprise have been paid the bondholders are the first class of security owners to share in any proceeds resulting from the sale of the corporate assets.

The *preferred stocks* of the corporation represent part ownership of the business, but they may or may not entitle the holders to vote in stockholders' meetings and thus play a part in the control of the business. Preferred stocks, like bonds, promise the owner a fixed rate of return on his funds, but this return does not have to be paid on the preferred stocks if the earnings of the enterprise are inadequate for the purpose, although it must be paid in any given year before any dividend payment can be made to the common stockholders. In the case of some preferred stocks, called "cumulative," any dividend payments which have been omitted in past periods must be made up and paid in full before any dividends can be paid to the common stockholders in later periods. Since the income of the preferred stockholders is considerably less secure than that of the bondholders, the rate of return offered on the preferred stock of any given corporation is likely to be higher than that promised on its bonds. Although the preferred stockholders are part owners of the corporation, once they have received their fixed rate of return they usually cannot share further in its earnings. However, some preferred stocks, called "participating," entitle their holders to a further share in corporate earnings as soon as the common stockholders have received dividends (in any period) as great as the fixed return promised to the preferred stockholders. Finally, preferred stockholders come after bondholders with respect to the right to share in any proceeds which result from the dissolution of an unsuccessful corporation and sale of its assets, but this frequently means that the preferred stockholders get little, if anything, once the claims of business creditors and bondholders have been satisfied.

The *common stocks* of the corporation, which also represent part ownership of the business, ordinarily give the owners the right to vote and otherwise take part in stockholders' meetings. However, the right to vote and to participate in controlling the affairs of the corporation means very little to most common stockholders in practice, as we shall see. Some corporations issue two or more classes of common stock and restrict the voting power to one class. This practice increases "owner-financing" and makes it possible to reduce the size of bond issues with their burdensome fixed charges without admitting more people even to the nominal right of controlling corporate affairs.

Common stocks do not promise any definite rate of return on the sums invested by the owners. The common stockholders may expect to receive dividends only if additional earnings are available after the bondholders and preferred stockholders have been paid their specified returns, and even then the corporation is not compelled to pay any dividends to the common stockholders. Instead, such additional earnings may be reinvested in the business, though this will usually have the effect of increasing the market value of the common stock per share. In some cases, if the corporation is highly successful, the returns to common stockholders on their investments may be very large. In other cases, common stockholders will receive only a small return or none at all. If the corporation fails, the common stockholders are the last to share in the proceeds which result from its dissolution and the sale of its assets. In practically all such cases, the common stockholders have no real hope of getting back even a part of their investments.

This discussion of corporate securities has had to do primarily with their status and relationships within given corporations, and care is necessary in applying it to the securities of different corporations. Within a given corporation, a bond is a safer and more conservative investment than a common or a preferred stock, but merely calling a security a bond does not automatically make it a safe and desirable investment. If a corporation, such as a holding company, has few assets except for its holdings of the stocks of other companies and proceeds to issue bonds on the basis of these stocks which it holds, the bonds may be very questionable investments indeed. Another company, with great physical assets and earning power, may be able to pay a high rate of dividends on its common stock through both good years and bad, and such common stocks may be preferable to the bonds of other companies.

Our references to the rights of various classes of corporate security holders may have given the impression that, if a corporation is unable to operate successfully, it is usually broken up and its assets are sold. As a rule, however, the assets of a corporation are worth more as part of a going concern than they would be likely to bring at a forced sale. Thus when a corporation is unsuccessful, it is more likely to be "reorganized" than liquidated. In the process of reorganization even the bondholders may have to make some sacrifice in order to keep the business in operation. That is, they may have to exchange large bonds for small ones or agree to take a lower rate of interest than that originally promised. Indeed, bondholders sometimes degenerate into stockholders in the process of reorganization and stockholders are often wiped out altogether. The result of reorganization is to scale down the annual fixed interest payments which the corporation must make and thus to improve its chances of operating successfully.

Limited Liability for Stockholders. The power of the corporation to acquire large amounts of capital funds is important, but this power is in

turn dependent to some extent upon other characteristics of this type of business enterprise, such as the limited liability enjoyed by the owners of the corporation, the relatively long life of the corporation, and the ease with which the shares of ownership in the business may be transferred. Partners and individual proprietors have unlimited liability, and their personal wealth and income are often taken for the satisfaction of business debts, but the stockholders of an unsuccessful corporation can ordinarily lose nothing beyond the amounts which they have invested in corporate securities. Some years ago stockholders in national banks had double liability; that is, if the enterprise failed, the stockholders could be called upon, if necessary, to contribute for the payment of its debts an additional amount equal to their original investment in the corporation. Although this is no longer true, state laws sometimes still impose double liability on the owners of certain corporations, but, for the most part, the personal wealth and income of stockholders are considered a thing apart from their investment in the business and cannot be tapped to pay the debts of the corporation.

Transferability of Ownership. Generally speaking, the shares of stock which represent ownership in the corporation are easy to acquire and easy to sell to someone else if the owner wishes to withdraw from the corporation. If the stocks of a corporation are listed and dealt in on one of the organized security exchanges, such as the New York Stock Exchange, they can be bought or sold on very short notice. Of course, it is not always possible to sell one's stocks for a total sum as large as that which was originally invested in them, but it is clearly much easier in general for the stockholders to withdraw their investments from the corporation than it is for the individual proprietor or the partners to accomplish the same result in connection with their enterprises.

Long Life. The corporation acts in business as if it were an individual itself, and it therefore has a life or a continuity of existence which is largely independent of the stockholders who own it. Unlike the individual proprietorship or partnership, the corporation's ability to continue in business is not usually affected by the death or withdrawal of individual owners. The banks which act as registrar and transfer agent in connection with the corporation's securities will record the changes in ownership which take place, but the business itself will go on functioning as if nothing had happened, except in cases where an important stockholder has also been active in the actual management of a business.

The Separation of Ownership and Control. The organization of a corporation involves somewhat more formality and red tape than that of an individual proprietorship or partnership, and it is also somewhat more expensive. The corporation may also be burdened with special regulations and with various kinds of taxes which do not apply to other forms of the business unit. On the whole, however, the corporation has very few disad-

vantages which may affect its efficiency and success as a type of business unit. If we are interested in discovering any major disadvantages of corporate operation, we must look for the effects of certain aspects of corporate operation on the stockholders, workers, and general public.

We have seen that the owners of the corporation (stockholders) cannot individually do business in the name of the corporation and that, even as a group, they do not actively manage the business of the corporation. Still more important is the fact that, in many corporations, the functions of ownership and control are widely separated in practice. The stockholders whose shares carry voting power are supposed to be both part owners of the business and the final source of authority in connection with the control of the corporation, but it is clear that in any case a group of individuals who owned a bare majority of the voting stock could control the corporation. Since a corporation may have a large investment made with funds borrowed from bondholders, and since it may have issued one or more classes of preferred stock and common stock without voting power, the owners of a majority of the voting stock may be able to control a very large amount of corporate wealth by means of a comparatively small investment of their own.

Moreover, it is not usually necessary for a group of individuals to own even a majority of the voting stock in order to control the corporation, for the average stockholder in practice is quite unlikely to try to exercise the rights inherent in his voting stock. He is not likely to be able to vote his stock in person, unless he can spare the time and money to attend the stockholders' meetings which are held annually for the purpose of electing directors of the corporation and transacting any other business which may come up. He can vote his stock by proxy, but the persons who offer to perform this service for him are ordinarily those who, as members of the board of directors, are already in control of the corporation's policies and appoint its officers. Since many people do turn their proxies over to the individuals who are already in control, the individual stockholder would find himself powerless even if he did attend the stockholders' meetings. Infrequently there may be a struggle for power between the incumbent corporate officials and another faction which seeks to displace them, and the small stockholder may find that his voting power has temporarily become a matter of some importance. In the ordinary course of affairs, however, the ordinary stockholder, even though he owns voting stock, is little more than a capitalist who furnishes the corporation with funds in the hope of receiving a return if the affairs of the business are managed successfully under the control of a small group of "insiders."

Of course, the control of corporations by small groups of individuals is not an evil in itself. In fact, such concentration of control might be regarded as necessary for the efficient operation of corporate businesses, but in

some cases it does lead to unfortunate practices. The insiders who control a corporation may provide unduly large salaries for themselves as officials of the organization. They may see to it that the corporation purchases machinery, equipment, supplies, or services at excessive prices from other companies which the insiders own or control. Such practices bring financial gain to the insiders while making sure that the corporation will have no large net income to distribute to the stockholders in general.

The insiders, or control group, may speculate in the securities of their own corporation and attempt to manipulate the prices of these securities so that their speculations will be successful. In some cases a small group of unscrupulous individuals in control of a corporation has withheld dividends from the stockholders and, through the spreading of rumors or false information, has made the stockholders think that the business was on the verge of failure when in reality it was operating quite successfully. Many stockholders, fearing that they would lose their entire investment in the business, would count themselves fortunate to be able to sell their shares for any price at all. These shares, of course, were snapped up by the controlling group, and eventually the business would begin to pay the large dividends it had been earning. In other cases, even though the business was operating at a loss, a group of insiders has been known to cause a corporation to pay good dividends until it could sell out its own holdings in the business at a favorable price and thus "get out from under." While there is nothing about the principles of economics to tell us that such practices are "bad," they are commonly so regarded, and attempts have been made by governmental regulations to eliminate or restrict many of them.

Overcapitalization. Shares of stock in a corporation are often issued with a certain nominal or par value specified on them which is supposed to represent the amount invested by the owner in the corporation. The total par value of all the securities of a corporation, called its *capitalization,* is supposed to equal the total investment in the business. When the capitalization of a corporation exceeds the actual amount of wealth or claims on wealth which it possesses, the corporation is said to be overcapitalized and its stock is said to be watered. Such overcapitalization may occur in a number of different ways.

Sometimes when a corporation is just starting out in business it may desire to purchase, say, a piece of land for a factory site. Preferring not to pay cash, it may offer to give the landowner stock in exchange for his land and, in order to interest him in the bargain, it may offer an amount of stock with a par value of $100,000 for land which would not sell for more than $50,000 or $75,000 in cash. Under some conditions this might be a desirable deal for the new corporation, but it would result in overcapitalization at least for the time being. In another case, overcapitalization may

be deliberately produced by the promoters of a corporation for their own benefit. They may sell bonds and preferred stocks to the general public to get the funds necessary for the establishment of the business, and issue the common stock to themselves either without payment or for a small consideration. In this way, since voting power is limited to the common stock, they may retain control over the business though they have little or no investment in it. If the business is successful, they can share in its profits or sell their common stock for a high price; if it is unsuccessful, they will have lost virtually nothing. Again, when a corporation has been in the process of promotion, prospective purchasers of preferred stock or even bonds have sometimes been offered a quantity of common stock without charge as a premium or inducement to purchase the other securities, much as some companies used to give away a "silver" teaspoon with the purchase of a pound of tea. In such cases, the purchasers of preferred stock or bonds appear to get something for nothing, and this may be the actual result if the business is successful, but for the present this common stock has no real assets of the corporation behind it.

Watered stock may also be issued after a corporation has been in operation for some time. For example, when a corporation has a chance to buy out a competing business, it may wish to use its securities instead of cash and, in its anxiety to acquire the assets of the other firm, it may offer the owners of the business an amount of stock which has a par value well in excess of the market value of the assets of the other company. Again, a corporation may be operating so successfully that its officials fear that its large earnings may attract competing companies into the same line of business or may lead to a governmental investigation of the corporation's affairs. By issuing large additional amounts of stock to the present stockholders without charge, the capitalization of the company may be so increased that the dividends will appear moderate on the basis of the par value of its stock. Dividends of $20 per share on stock with a total par value of $1,000,000 could thus be converted into dividends of only $5 per share on stock with a par value of $4,000,000.

Thus watered stock may be used as a device to conceal the earnings which a company is making on the actual investment in the business. Moreover, some security purchasers may actually be misled by the par value of securities and may think that these values actually represent the value of the assets which the company is holding behind the stocks. Finally, in the case of a regulated monopoly, such as a public utility company, or a large combination of business units, a large capitalization which consists in part of watered stock may be made the basis of a claim that the company needs to charge higher rates or prices in order to afford its owners a reasonable rate of return on their securities. Under strictly competitive conditions, of course, a company which is overcapitalized cannot charge

higher prices than one which is not. It might wish to charge higher prices, in order to be able to pay good dividends on its stocks, but if it tried to do so its competitors with their lower prices would get all the business. The abuses of overcapitalization have been very serious at times, and many laws have been passed to regulate the activities of corporations in this matter. Some corporations have avoided the question of overcapitalization by issuing stocks with no par value. In such cases the only value which is likely to be assigned to the corporation's assets by the public is that equal to the total market value of its securities.

Stock Dividends. Stock dividends occur when a corporation issues additional shares of stock without charge to the present stockholders on the basis of their holdings of the company's stock. In some cases such stock dividends may be entirely justified and desirable. If a company is operating successfully with large earnings and wishes to expand its business, it can pay out these earnings as dividends and then try to sell additional securities in order to get the funds necessary for expansion. It is much more simple and usually more economical, however, to reinvest the earnings in the business and issue more shares of stock to the present owners to represent the increase in assets which takes place. If any stockholders desire their dividends in cash, they can simply sell the additional shares of stock which they receive. A similar result would be more difficult for the stockholders to accomplish if the corporation reinvested its earnings without issuing additional stock, although in such an event the market value of each share of the stock already in existence would tend to increase.

In other cases, in an effort to keep stockholders contented, a corporation may issue stock dividends even though it has no actual earnings to pay out or reinvest. If the stock has par value, the result of this process is overcapitalization, since capitalization has been increased without any increase in corporate assets. If the corporation's stock has no par value, it cannot become overcapitalized by the issuance of stock dividends for, even though its assets have not increased, the only effect of the stock dividends will be to cut the ownership of its present assets into thinner slices. However, the value of corporate assets may be written up by corporations which issue stock with no par value.

Relations with Workers and Customers. Many of the abuses which have developed in connection with the corporation have depended upon the almost complete ignorance of the stockholders with regard to the affairs of their corporations. Although modern laws require corporations to make their financial statements available to stockholders, most of such statements are so brief and technical that many stockholders know almost nothing about the businesses in which they own stock. In fact, even in modern times, individuals may act upon the advice of friends and buy Air Reduction, Alaska Juneau, Alleghany, or some other stock without even knowing

the kind of product which the corporation produces. These are extreme cases; most stockholders today have a general knowledge of their corporations' products and financial standing, but they usually have no contact at all with the workers whom they indirectly employ or the customers who are served by their businesses.

On the other hand, individual proprietors or partners, unless their enterprises are very large, are likely to be in fairly close touch with their employees and their customers. Although they may no longer work side by side with their employees, they know rather accurately the wages paid, the hours worked, and the conditions under which the work is done. The stockholder in a large corporation, satisfied with receiving dividends regularly on his investment, is not likely to know or care whether the wages paid by the corporation are high or low, whether the hours worked by employees are short or long, or whether the working conditions are satisfactory or miserable. In similar fashion, he may know or care little concerning the prices charged by his corporation, the service which it renders, or the quality of its relations with the public. The fact is that, in dealing with the corporation, corporate employees and customers are up against a blank wall and cannot reach the actual owners in trying to secure satisfactory hours, wages, and working conditions or satisfactory prices and service. These conditions may have important consequences in terms of labor efficiency, the prevalence of labor unrest and disputes, and the existence of unsatisfactory relations with the public. Of course, modern corporations often make an effort to remedy these situations, but the problems implicit in the impersonal character of the corporation remain troublesome.

Corporate Domination of Economic Activity. As our last point in connection with the corporation, we note that many people today fear that a few large corporations will come to dominate the entire economic life of the country, if indeed they do not already do so. We have seen that, in 1937, corporations produced some 60 to 65 per cent of all income in the United States, and they also owned almost 80 per cent of all business wealth. Moreover, the ownership of corporate wealth was concentrated in the hands of a small number of large corporations. In 1937 the 394 largest corporations owned about 45 per cent of all corporate wealth, though they made up less than one tenth of 1 per cent of the total number of corporations.[1] We have some 30 "billion dollar" corporations in the United States. Each of the two largest of these corporations (the Metropolitan Life Insurance Company and the American Telephone and Telegraph Company) has larger total assets than any of 38 of our states.[2] In 1936, the 200 largest nonfinancial corporations had total assets of over 75 billion dollars and the

[1] Temporary National Economic Committee, *Final Statement of Senator Joseph C. O'Mahoney*, p. 7.

[2] *Ibid.*, p. 5.

50 largest financial corporations had total assets of some 35 billion dollars, a total of 110 billion dollars or close to one third of the national wealth.[3] Corporate income is concentrated in similar fashion. In 1937, 248 corporations received 40 per cent of the total net income of all corporations, though they made up only one tenth of 1 per cent of the total number of corporations with incomes.

It is sometimes argued that this concentration of corporate wealth and income is not very significant, since the large corporations themselves are owned by very large numbers of stockholders. However, as we have seen, most corporations are *controlled* by a relatively small number of individuals, and even corporate ownership is not nearly so widespread as is popularly supposed. For example, it has been estimated that one half of all dividends paid in the United States are received by stockholders who make up less than 1 per cent of all American stockholders.[4] There may also be a tendency in the direction of greater rather than less concentration of corporate wealth and income, for the large corporations seem to do the greater part of corporate saving. In 1937, the corporations with assets of $1,000,000 or more made up only 5 per cent of a group of nonfinancial corporations studied, but they made 88 per cent of the savings of all these corporations.[5] When we realize that the large corporations of the country are more or less closely interconnected through interlocking directorates and interest groups, and that the influential individuals in control of these corporations often play an important part in controlling the government and public opinion of the country, we can understand why some people fear the control of economic life by the large corporations.

THE COOPERATIVE ENTERPRISE

A cooperative enterprise is *an association of individuals for the carrying on of some type of economic activity for the mutual benefit of the members.* This definition is so broad that it seems capable of including almost any type of business enterprise, such as partnerships or even corporations. As a matter of fact, the chief difference between the cooperative enterprise and other types of enterprises is not so much in the form of organization as in aims and methods. Cooperative enterprises may actually be organized in the corporate form, but they do not usually seek to make profits for their members by dealing with the general public. They ordinarily carry on most of their business dealings with their members and for the benefit of these members. If a cooperative enterprise is organized as a corporation, the various members are likely to have equal voting powers without regard to

[3] National Resources Committee, *The Structure of the American Economy.* Washington, D. C.: Government Printing Office, 1940, pp. 274–276 and 298.

[4] Temporary National Economic Committee, *Final Statement of Senator Joseph C. O'Mahoney*, p. 8.

[5] *Ibid.*, p. 9.

the amounts of stock which they own individually, and they are likely to share in the earnings or savings realized by the business on the basis of the amount of their dealings with the cooperative enterprise and not on the basis of stock ownership.

Consumers' Cooperatives. Consumers' cooperatives are a well-known type of cooperative enterprise, although they are not of great importance in the United States. They are usually started for the purpose of saving for their members as much as possible of the difference between the prices which ordinary retailers pay for goods and the prices for which they sell, but they may go on later to operate wholesaling or even industrial establishments for the benefit of the members as consumers. The original funds for starting such a cooperative may be obtained from membership fees, loans from members, or the sale of stock, but the funds invested in the enterprise draw only common interest and are not made the basis for sharing the savings or earnings realized by the enterprise.

A consumers' cooperative deals directly with wholesalers or form producers, getting whatever it can in the way of low prices or discounts on the commodities which it buys. These products it offers for sale in the retail store or stores which it operates, and it usually charges for its products just about the full retail prices which prevail in other stores. Though operated for the benefit of members, the cooperative stores often sell both to members and nonmembers. The savings or earnings which the enterprise secures through all these purchases and sales are used from time to time for three purposes. A part may be devoted to educational work and another may go to increase the capital funds of the enterprise, but the largest part is likely to be distributed to the members and forms the benefit which they derive from their membership in the enterprise. These "patronage dividends" are divided among the members on the basis of the total amount of purchases made by each member at the cooperative store during the business period in question.

Although the cooperative store sells to nonmembers, these persons are not often permitted to share in the earnings of the enterprise; but the earnings realized on sales to nonmembers are not ordinarily distributed as patronage dividends to the members, for this would make the business to some extent a profit-making rather than a cooperative enterprise. Sometimes the earnings on sales to nonmembers are allowed to accumulate to their credit in the treasury of the cooperative until such time as they become large enough to permit the nonmembers to purchase a membership in the cooperative. Membership is open to almost anyone who wishes to take part in the affairs of the cooperative. The control of these affairs is democratic in character, since each member ordinarily has one vote without regard to other considerations.

Credit Cooperatives. In associations known as credit cooperatives or credit unions, the members of the group place in the treasury of the cooperative any amounts of money which they wish to save. Individual members may then borrow funds from the cooperative for their personal use in accordance with certain regulations. The rates of interest charged to borrowing members are generally higher than the rates which could be secured for these funds in ordinary investments but they are lower than the borrowing members could secure if they tried to obtain funds for their personal use from a private credit agency. Thus the borrowing members gain by borrowing at a rate lower than those charged by ordinary personal finance organizations, while the depositing members who do not borrow get more interest on their savings than they could obtain in the usual run of investments and, in addition, have a source from which they too can borrow if necessary. Credit cooperatives also exist in the form of building and loan associations and kindred organizations.

Producers' Cooperatives. When a number of workers get together, pool their financial resources, hire or buy the necessary facilities for production, elect or employ a manager, and operate an industrial establishment, the result is a producers' cooperative. Although an enterprise of this kind is undoubtedly run for the benefit of the cooperative group, its members benefit by selling goods to the general public at a profit (if possible) as well as by earning their wages cooperatively. Thus the enterprise really differs but little from other forms of business organization, except that the workers are the enterprisers and receive their income as a residual amount after paying the other expenses of the business, instead of receiving wages from an enterpriser.

This type of cooperative has never gained a secure foothold in the United States and has not been very important elsewhere. The closest approach to producers' cooperatives in the United States is found in our various marketing cooperatives, which handle and sell for their members such things as grain, livestock, milk, fruits, and vegetables.[6] The members of such cooperatives remain independent in so far as their productive operations are concerned, and cooperate only in selling their products or in purchasing various materials and supplies. Many of these cooperatives are purely local in scope, though in some cases they cover a very wide area. In general they aim to secure for the members better prices than the members could obtain as individuals. This objective may be attained in the case of the larger cooperatives or federations of cooperatives through some combination of factors such as better storage of products, better grading of products, large-scale sales, education of growers as to the best types of products and methods of preparing products for the market, thorough knowledge of markets, demand creation through advertising, adjustment

[6] These marketing cooperatives are usually called producers' cooperatives in this country.

of freight claims for members, and attempts to secure correction of abuses in the markets for the products. In general, marketing cooperatives have been more successful than other types in this country, though the cooperative movement as a whole has never attained here the proportions which it has reached elsewhere.

THE GOVERNMENTAL ENTERPRISE

Types of Governmental Enterprise. Finally we must note that even in our capitalistic system certain phases of economic activity are carried on not by private individuals or limited groups of individuals but by governmental units which presumably act for all the people within their jurisdiction. Some governmental activities are undertaken to furnish people with services—such as national defense or the maintenance of law and order—which they could hardly provide for themselves privately. Other lines of production, such as the furnishing of gas, electricity, or water to communities, can well be carried on by private enterprises in some cases, but they may be taken over by governmental units with the hope of giving consumers more satisfactory service or lower rates. Instead of being left in private hands the sale of liquors may become a governmental activity, partly because it seems desirable to set some effective limits on it. Hospitals and institutions for dependents and defectives may be operated publicly for reasons of economy or because, although they could be run privately, private citizens might not have sufficient incentives to do so.

In general, under capitalism, it is assumed that the various governmental units will not operate business enterprises in direct competition with private enterprises in given fields. The activities which are carried on by governments are often reserved entirely to them, and fields which are left in general to private enterprise are seldom invaded by governmental units. Thus, in the great depression after 1929, when millions of persons were unemployed and the federal government decided to furnish them with employment, the unemployed were in general not set to work by the government to produce clothes, shoes, food, and other necessities of life, although these things were badly needed by many people. Instead, the government undertook to harness the tides of Maine and to cut a canal across Florida, or set people to work at raking leaves or constructing public buildings.

However, there are some exceptions to the general principle stated above. Educational activities are carried on both by governmental units and by privately owned educational institutions. The federal government, through the Tennessee Valley Authority, has entered into quite direct competition with some privately owned public utility companies. During World War II the federal government constructed or took over and operated numerous plants in industries which were for the most part in private

hands. Between 1940 and August, 1944, the construction of over 20 billion dollars' worth of manufacturing facilities had been provided for in industries producing for war purposes, and some 78 per cent of this amount, or 15.9 billion dollars, was financed by the federal government.[7] On the whole, though, governmental units under capitalism respect the boundaries of the fields which are traditionally left to private enterprise.

Some governmental enterprises, such as the salt or tobacco monopolies in some countries or liquor monopolies carried on by some of our state governments, are the source of considerable revenue for the governments in question. The governmental unit needs a certain total amount of revenue, and the more it can derive from governmental enterprises the less it needs to extract from the citizens through taxes. Such public enterprises are often run on the basis of price principles similar to those of private monopolies. Other governmental industries, such as the post office in the United States, make charges for their services but do not attempt to derive the greatest possible revenue from their activities. The fact that our government's postal operations are seldom conducted at a profit is considered by many people as proof positive of the inefficiency of public enterprise, but there is little doubt that these operations could be conducted at a considerable profit if such a result were desired. Instead, it is desired to give our citizens the benefit of nation-wide, efficient communication as nearly as possible on a cost basis. Sometimes the post-office system just makes ends meet, while at other times a profit is made or a loss is incurred. Finally, some governmental activities are carried on without any thought of deriving from the sale of the services an income which will cover the costs of furnishing them. The public school system is an example. Anyone may partake of the service without charge, the costs being paid out of taxation. A childless millionaire will contribute heavily to the support of the school system through taxation, while the poor man with a large family may send all his children to school even though he pays only a very small amount of tax.

Forms of Business Enterprise and Economic Systems. The discussion of the forms of the productive unit in this chapter applies almost entirely to a capitalistic system such as our own. Under socialism and communism, virtually all departments of economic activity would be owned and operated by society as a whole, though under socialism some individual enterprisers and cooperatives might be tolerated. Clearly it would be useless to discuss the merits of the corporation or partnership as forms of business enterprise in connection with an economic system in which only governmental enterprises prevail. Under fascism, the various private *forms* of organization were for the most part retained, but the *control* of these productive units

[7] K. C. Stokes, "Financial Performance of Large Corporations," *Survey of Current Business*, August, 1945, p. 9.

rested almost completely in the hands of the government or leaders of the ruling party. Thus the present chapter is one which would be of slight, if any, importance for a book on the principles of economics for use in any type of economic system other than capitalism.

QUESTIONS AND PROBLEMS

1. "Single proprietorships are very numerous but are not very important in terms of workers employed and goods produced." Why?

2. "Partnerships may enjoy several advantages in business as compared with single proprietorships, but almost none in comparison with the corporation." Explain.

3. "The corporation is strong where partnerships and single proprietorships are weak." Show whether you agree.

4. Distinguish between bonds, preferred stocks, and common stocks.

5. "The levying of a heavy tax upon profits would seriously impair the value of both stocks and bonds." Do you agree? Explain.

6. Show whether you would rather be a common stockholder, a preferred stockholder, or a bondholder in a corporation (1) when the business is prosperous, (2) in the event of liquidation.

7. "The stockholders of a corporation ordinarily manage and control the business in addition to owning it." Show whether you agree.

8. "The capital-raising power of the corporation depends to an important extent on certain other advantages of this type of business unit." Explain.

9. What are some of the evils which may result from the control of a corporation by a small group of insiders?

10. "Overcapitalization is to be condemned because it enables a corporation to charge higher prices for its products than it could otherwise obtain." Do you agree? Explain.

11. How may a corporation become overcapitalized? Explain.

12. "The issuance of stock dividends by a corporation may or may not be justifiable." Explain.

13. "A corporation is more likely to experience labor troubles than is a single proprietorship or partnership." Show whether you agree.

14. Why do some people fear corporate domination of all economic activity in the United States? Explain.

15. "One possible disadvantage of the corporation is found in the fact that it is a more impersonal organization than the single proprietorship or partnership." Explain.

16. "Cooperative enterprises differ from corporations in aims and methods rather than in form." Do you agree? Explain.

17. Why have cooperative enterprises never attained great importance in the United States?

18. "Governmental enterprises may be undertaken for any one of a variety of reasons." Explain.

19. "Governmental enterprises usually compete directly with private enterprises in given fields of production." Show whether you agree.

20. "The fact that governmental enterprises seldom make a profit indicates that such enterprises are usually managed inefficiently." Do you agree? Explain.

REFERENCES FOR FURTHER READING

Berle, A. A., and Means, G., *The Modern Corporation and Private Property.* New York: The Macmillan Company, 1932, Book I, Chaps. 1–6, and Book IV, Chaps. 1–4.

Big Business—Its Growth and Its Place. New York: Twentieth Century Fund, 1937, Chaps. 1, 2, 4, 5, 8.

Black, J. D., *Production Economics.* New York: Henry Holt and Company, 1926, Part IV, Chap. 18.

Dewing, A. S., *The Financial Policy of Corporations.* New York: The Ronald Press Company, 1934, Book I, Chaps. 1–6.

Dixon, R. A., *Economic Institutions and Cultural Change.* New York: McGraw-Hill Book Company, Inc., 1941, Chaps. 16–19.

Ripley, W. Z., *Main Street and Wall Street.* Boston: Little, Brown and Company, 1927, Chaps. 1–11.

Robinson, E. A. G., *The Structure of Competitive Industry.* New York: Harcourt, Brace and Company, 1932, Chap. 4.

IX

Risk in Economic Activity

Regardless of the form in which the productive unit is organized or of the size to which it grows, economic activity is always filled with a variety of risks. Many of these risks are the special concern of business enterprisers—those persons who assume ultimate responsibility for the conduct of business enterprises. In all the fields of activity which we include under the term "production"—in form production, transportation, marketing, finance, storage, and even in the rendering of personal services—those in charge of business enterprises are constantly faced with the necessity of making decisions and, in the case of each decision, there exists a degree of uncertainty as to the proper course to be followed and a risk that the decision may produce results which are unfavorable to the enterprise. In addition to the risks involved in the economic decisions which must be made, there are many factors these decisions cannot control which may produce favorable or unfavorable results for any enterprise.

THE NATURE OF RISK

Risks of Roundabout Production. It is true that economic activity has always been risky, but it should be noted that many of the risks to which business enterprises are exposed today have developed to their present importance in comparatively recent times. In modern economic activity, production is almost always indirect or roundabout; that is, it is carried on by means of large quantities of capital, in addition to labor and land. Much of the capital required for indirect production is fixed in character. In addition to being expensive, fixed capital goods are durable and capable of performing their functions over a period of some years before wearing out. Once an enterpriser has made a large investment in fixed capital goods, such as machines and buildings, he may find that their productiveness exceeds his fondest expectations and that they will, for a time at least, furnish him with an income which is much more than adequate to cover the interest which he may have agreed to pay in order to obtain the necessary funds for investment. In other cases enterprisers discover that they

have acquired fixed capital goods which, in view of market and other conditions affecting the industry, will not earn their keep. Such mistakes can ordinarily be corrected only over a considerable period of time. Moreover, even if the judgment of business enterprisers in selecting fixed capital goods is as accurate as possible at the time, the discovery of new machines and methods of production may soon render these fixed capital goods obsolete from an economic point of view long before their physical usefulness is at an end. Such developments often cause heavy losses to business enterprises. Many years ago, when land and labor in most lines of production were assisted only by a few simple hand tools, this type of risk clearly was not of any great significance to business enterprisers.

Risks Resulting from Large-Scale Production. Because of a number of factors which were discussed in Chapter VII, large-scale productive units are the rule today in many fields of economic activity. However, in turning out a large volume of production, enterprisers are often unable to know just who will buy the products, when they will be sold, or what price they will be able to command. Such enterprisers are said to produce for a future market and in anticipation of demand. Their position is decidedly risky. The individual enterpriser is anxious, of course, to turn out the volume of product which will furnish most favorable financial results for his firm. It would be difficult enough for him to reach an accurate decision on this point if he were the only enterpriser in the field, but when, as is often the case, he is quite uncertain as to what others in the same line of production or in related industries are doing, his problems become even more troublesome.

Again, the enterpriser can never be certain as to what will happen to the demand for his product. The tastes of consumers are subject to frequent change, and the enterpriser may find at any time that the former consumers of his wares are inclined to satisfy their wants by consuming other goods which may be substituted for his, or other varieties of the same product. Business conditions may change for the worse during the period of production, reducing the incomes of large classes of consumers and decreasing the demand for the goods of various enterprisers. In other cases an improvement in general business conditions or a favorable change in the tastes of consumers may bring unexpected gains to an enterpriser. The risk that he will not be able to get for his product, when it finally reaches the market, a price as great as he expected when production was undertaken is one of the most troublesome risks which the business enterpriser must face.

Even if the enterpriser is fortunate enough to be able to sell his goods in advance, so that he knows what price he will receive for his products, he is by no means free of risk. His raw materials may advance in price, his machinery may break down, his workers may go out on strike, or any one of a large number of factors may operate to increase his costs so that

the price to be received for the products will not be enough to furnish him with the expected amount of net income. Many years ago, when production took place in the individual workshop on a small scale and largely on order, the risks involved in marketing products were relatively small. To be sure, the enterpriser's shop might burn down, his customer might die before the goods which had been ordered had been produced, delivered, and paid for, or an employee might turn out to be dishonest, but the risk involved in producing on a large scale in anticipation of demand had not yet arisen.

Risks Resulting from Specialization. Under a system of nonspecialized production where the members of each family attempted to engage in most types of production necessary for their economic welfare, the failure of any one or two types of productive activity, while unfortunate, might not seriously reduce the standard of living enjoyed by the members of the family. Under specialized production, on the other hand, the owners and workers in a particular industry concentrate their eggs in one basket, or a small number of baskets, so that if anything happens to depress this one industry or impair its successful operation the results are likely to be most severe for owners and workers alike, and perhaps also for people in other lines of production which are in some way dependent upon this one. Even though all the enterprisers in a particular industry plan their affairs with the utmost care, we must remember that, in our type of economic system, there is no central authority to coordinate the productive activities of the various industries and it is relatively easy for maladjustments to occur between industries. Specialization leads to great interdependence among the parts of the economic machine and intensifies the risks of business.

Other Risks. In addition to the risks which have been described, there are others which, although of considerable importance in our modern economic system, are not peculiar to that system but have long troubled the business enterpriser. The enterpriser in agriculture is at the mercy of weather conditions. Windstorms or hailstorms may damage his crops, duststorms may ruin his land, and too much or too little heat or moisture during the growing season may cause the failure of his crops. Even favorable weather conditions may not benefit the farmer financially, for large crops often sell for a very low price per unit of product so that growers may receive a smaller total sum for a large crop than for a moderate or small one. These factors which influence agricultural prices and production also affect the manufacturing, transportation, storage, and other industries which depend upon agriculture either for raw materials or for part of their volume of business. Manufacturing establishments are not so directly concerned with weather conditions, but they, like establishments in other lines of production, may be damaged by fires, floods, and other

natural phenomena. In any field of production, valued employees or officials may die or be lured away by other concerns. Trusted employees or officials may turn out to be dishonest and almost any productive establishment may be robbed or burglarized. And there is the constant danger that productive facilities may break down or labor troubles crop up.

Risks of Landowners and Capitalists. We have been discussing risks in economic activity as if they consisted almost entirely of phenomena which must be dealt with in some way or other by business enterprisers. This really comes close to being the case, for business enterprisers, as persons who assume ultimate responsibility for the conduct of business enterprises, are the residual risk bearers in our economic system. They seek, as we shall see, to eliminate or reduce risks wherever possible or transfer them to the somewhat broader shoulders of professional risk bearers, but the risks which remain must be borne as best they can be.

Landowners, as such, have comparatively few risks to bear. To be sure, agricultural land may be damaged by misuse or through erosion by wind or water, and land which is used for site purposes may lose value as the town or city grows away from it in some unexpected fashion, but in general the holders of land in the midst of a growing population are not in a very risky position. There is also the chance that tenants may be unable to pay their rents, but this risk is supposed to be allowed for to some extent under competitive conditions in the contract rents which are charged to tenants.

Capitalists, like landowners, are not among the leading risk bearers in economic activity. There are, as we have seen, many things which can happen to damage or destroy capital goods as physical objects of wealth or to impair their value from an economic point of view. There are also many factors which may cause a business enterprise to operate so unsuccessfully that it cannot pay interest or repay capital funds when these payments are due. However, the gross interest rates which are received by capitalists when they entrust their funds to one industry or another are supposed under competitive conditions to include a payment which compensates for the varying degree of risk involved in the different fields of economic activity. In particular cases, capitalists may lose their all, but in the long run and for the economic system as a whole their risks are compensated for quite adequately by the gross rates of interest which they receive in their various investments of capital funds.

Risks of Workers and Business Enterprisers. The workers in economic activity have relatively heavy risks to bear. Modern mechanized production affords almost limitless opportunities for industrial accidents, and millions of such accidents, of varying degrees of seriousness, occur in each year. In other millions of cases, workers contract different kinds of occupational diseases or ailments as a result of the unsatisfactory conditions under which they work. Colds, grippe, or pneumonia may incapacitate

workers who are subjected to sudden and drastic changes in temperature, improper ventilation, or the necessity of breathing excessively moist air. Tuberculosis, silicosis, anthracosis and other lung diseases may strike down workers who perform their functions in an atmosphere charged with minute particles of materials on which or around which they work. Employees who work with lead, radium, or other metallic elements are in danger of metallic poisoning, and so on.

Even if workers keep their health, they run the risk of becoming unemployed; that is, of being able and willing to work but unable to find employment. Many workers are thrown out of employment at least for a time because of the seasonal nature of the industries in which they work. Most industries show some seasonal variation in the rate at which they operate, and some, such as automobile production, building construction, coal mining, and clothing production, are highly seasonal in character. In the dull season in such industries, many workers are laid off, and it has sometimes been estimated that, on the average, as many as 5 per cent of all our workers are kept idle because of seasonal fluctuations in production. Even more important is the amount of unemployment which results from cyclical fluctuations in business activity. Although some unemployment seems to exist even in the best of times, the volume of unemployment often increases sharply as production falls off in depression years. In the worst of the depression years following 1929, for example, the number of unemployed was sometimes estimated at about fifteen millions, or close to one third of the number of those who had been gainfully employed in 1929.

Technological changes, as we shall see in detail later, tend to create as much employment in the long run as they destroy, but at any one time there is likely to be a considerable amount of technological unemployment. While the workers displaced in this way do tend to find some kind of employment sooner or later, the processes of re-employment may require considerable time. The new jobs obtained by the workers may not be nearly as pleasant or remunerative as the old ones, for specialized skills, developed and trained over a long period of time, may be almost or quite valueless in the new occupations. Finally, whether or not they suffer through periods of unemployment in the meantime, workers may eventually reach the day when, from the point of view of industry, they are too old to hold down their present jobs or secure new ones, though their needs for consumers' goods may continue virtually undiminished for many years to come. In recent decades many industries have come to depend more and more on young workers and have gradually lowered the maximum age at which they will hire new employees. This premature obsolescence of the human machine means permanent "unemployment" for the persons affected and it therefore constitutes a major risk for these workers.

These risks which are faced by workers are compensated for to some

extent under long-run competitive conditions by the payment of different rates of wages in occupations which involve varying degrees of risk. If an occupation subjects the worker to great danger of accidental injury or occupational disease, the wage which is paid ordinarily should be greater than that which prevails in other safer occupations which are open to the same class of worker. If production and employment are highly seasonal in a certain industry, the daily wage tends to be well in excess of that received by other workers of the same grade whose jobs are steady the year around. If a worker has a type of employment in which he can be relatively sure of functioning until he reaches an advanced age, he may expect his wage to be lower than that received by a worker in some other occupation in which his economic career may end at forty or forty-five. However, workers are likely to underestimate the risks involved in the occupations which they undertake and their bargaining power may not in many cases be sufficiently strong to insure adequate differentials in pay for the more risky occupations. Moreover, the risks which workers face result in personal hardship and suffering when they turn out unfavorably, in addition to loss of wealth and income, and it is not always possible to provide adequately for injury, loss of health, worry, and other personal factors through the simple medium of a money payment. As we shall see later, the risks of the workers can often be reduced significantly by careful planning within the business enterprise, and it is possible to shift a part of their financial burden to others through some form of insurance.

The risks faced by the workers are personal risks, and are not to any significant extent risks of the business enterprise as such. When, as sometimes happens, the workers are left to bear their risks as best they can, the financial losses which they may be called upon to sustain are very heavy in proportion to their ability to bear such losses. The financial losses which business enterprisers may face, when their risks result unfavorably, are large in absolute amount but no larger, and are perhaps smaller, in proportion to their ability to bear such losses, than the possible losses of workers. However, the risks of the business enterprise as an economic unit are risks which the business enterprisers must handle in some way and, if they do not wish to bear these risks, they must seek to reduce them or transfer their burden to someone else.

Economic Risks and Economic Systems. From the point of view of the economic system as a whole, the risks which business enterprisers face as the result of their own economic decisions are of little significance in the long run, for the favorable outcomes tend to cancel the unfavorable ones. However, in a capitalistic economic system they are of great significance to the individual enterprises which face them. In other economic systems in which economic activities are operated by society as a whole and productive wealth is socially owned, the problem of business risk is not impor-

tant. The system is large enough so that losses from natural causes can be accurately estimated and provided for. The risks of profit or loss as a result of economic decisions and commitments in the various lines of production are of no significance, since there are no private enterprisers to make such profits or sustain losses. The economy as a whole is the enterpriser and considerations of *financial* profit or loss do not mean anything to a whole economy. However, the risks of workers, as personal risks, would need to be provided for in such a system.

INSURANCE

The Reduction of Risks. When an individual is faced with risks which he does not want to bear, the first logical step would seem to be an attempt to reduce them. If an automobile owner is greatly worried about the risks involved in operating his vehicle, he may try to safeguard himself by keeping his car always in the best possible mechanical condition, by obeying all the rules and regulations of the road, by adapting his speed to the existing driving conditions and his own condition, and by extending every possible courtesy to the drivers of other vehicles. If a business enterpriser is concerned about the risk of fire, he may attempt to make his establishment as fireproof as possible in its construction, install a sprinkler system, and prevent the accumulation of waste materials. Again, to protect himself against dishonesty on the part of his employees or other people, he may select his employees with great care, see to it that employees with small incomes are not placed in positions where they must handle large sums of the firm's money, install burglar alarms, and arrange to have the firm's receipts collected and its pay rolls delivered by other concerns which specialize in the transportation of funds.

The enterpriser who is impressed with the need of doing something about the risks faced by his workers may conduct educational, safety-first campaigns in the factory, screen or otherwise guard the moving parts of machinery to prevent accidents, have dangerous work performed by machinery as far as possible, see to it that conditions of light, heat, and ventilation are appropriate for the workers' health, and furnish workers with equipment to protect them from the risks encountered in breathing dust-laden atmosphere or handling dangerous industrial materials. He may protect workers against seasonal unemployment by taking on side lines of production which have their busy season when the main line of production has its dull season or by guaranteeing workers a certain number of weeks' work or pay in each year. He may ease the burden of technological changes by introducing such changes gradually, rather than suddenly, and by carefully selecting the workers who are to be displaced by such changes, and so on.

The Principles of Insurance. In spite of all that can be done to reduce the various risks found in economic activity, they cannot ordinarily be entirely eliminated, and the risks which remain may be too large for the individual to bear. In such a situation he may be able to transfer the burden of the risks through insurance. Therefore we should examine here several of the principles of insurance; one such principle is that, while the occurrence or nonoccurrence of a given event may be entirely unpredictable in any individual case, its probable incidence may be accurately estimated if a large number of such cases can be grouped together. Thus it may be regarded as almost completely uncertain whether any given individual will die in a certain year or whether any given factory building will be damaged or destroyed by fire or other hazards. But insurance companies, by gathering together large numbers of such risks, can predict with great accuracy, not which people will die or which factories will burn, but how many losses of each kind will occur in a certain period and what total sums will need to be paid to compensate for these losses. These predictions of insurance companies are based on careful records of the losses which have been sustained in the various classes of risks in the past.

Through insurance, uncertainty for the individual is converted into relative certainty for the group. The insurance company estimates the losses which will be sustained on account of a certain risk faced by a large group of persons or firms and spreads this loss over the group by charging each insured person or firm a certain premium which is related to the exact nature and amount of the risk which he faces. Then the funds collected in this manner, or most of them, are used to compensate those members of the group for whom the risk in question results unfavorably within the given period. Thus the insured individual or firm substitutes a small certain loss (the necessary insurance premium) for the much larger uncertain loss which would occur if the outcome of the risk were unfavorable to him. As a factory owner he may pay a premium for fire insurance year after year and never have a fire, but he is at least sure that, if his factory burns down, the loss will fall on the insured group as a whole and that he will never have to bear such a loss unassisted.

Insurance results in a reduction in uncertainty. It does not necessarily reduce the risk itself. A factory or store which is insured may be just as likely to burn down as one which is not. In fact some cynics suggest that a business establishment which is insured is more likely to burn than one which is not. In practice, however, insurance may result in a reduction in risk as well as in uncertainty. The insurance companies, with concern for the interests of insured persons and firms as well as for themselves, often conduct educational campaigns designed to reduce the risks that are insured against. Moreover, business enterprisers are likely to take steps to reduce fire hazards or the danger of accidents to their employees when they learn

that the premiums they are required to pay for insurance are based on the conditions that they maintain in their establishments.

There are some other principles of insurance which may be worth mentioning. A risk, to be insurable, should not be one which can result unfavorably for all, most, or even large numbers of the insured persons or firms in any given period. That is, unless the various individual risks are independent of each other, a loss sustained by any one of the insured may lead to similar losses by others. If a fire insurance company insured only a number of like houses or factories in a certain area of a city, so that a large fire might wipe out all or most of the insured properties, its position would be unsafe. Deposit insurance for banks and unemployment insurance for workers are sometimes suspect because of this principle. The failure of one bank may start fatal runs on other banks with the result that a large number of institutions may suffer losses within a comparatively short period of time. Clearly, too, a considerable part of our working population may be unable to find employment at a given time in a period of depression, and when workers are thrown out of employment in one industry their loss of income and inability to buy the products of other industries may lead to unemployment in other branches of production. Thus bank failures and unemployment are apparently cumulative, and "insurance" against these hazards is likely to be undertaken only by public authority, if at all.

An insurable risk should also be of such a kind that it is readily possible to determine when a loss has occurred as a result of the unfavorable outcome of the risk. It is difficult to pretend that you have lost an arm, but much less difficult to simulate sickness or unemployment. Thus the risk of industrial accidents comes nearer to being an insurable risk than the risks of industrial illness or unemployment. Finally, it is clear from our previous discussion that a risk is scarcely insurable if it is one which affects only a few persons, or if no records of previous experience as to losses due to the risk are available.

Business enterprisers are able to protect themselves through insurance against losses that result from the many risks which qualify as insurable, but there are other risks which still remain to trouble the enterpriser. It is difficult, if not impossible, to insure against the risk that fixed capital goods will become obsolete long before their physical usefulness has disappeared. Insurance companies will not guarantee the enterpriser that he will get, when his goods reach the market, the price that he expected to get for them when he started production, nor will they protect him against an increase in costs if he has contracted to deliver goods in the future at a fixed price. Even the risk of labor troubles must be borne by the enterpriser. However, in some cases, by means of a practice known as "hedging,"

enterprisers are able to protect themselves against certain risks of loss due to price changes. This device will be discussed later in the chapter.

Workers often do not have funds adequate to secure protection for themselves against the risks which they face in economic activity, and the nature of some of these risks is such that private companies might not be willing to offer insurance against them in view of the principles previously discussed. In this country,. however, almost all of the states have workmen's compensation laws which require employers to provide benefits during a certain period of time to workmen who have been injured on the job, and to pay a certain lump sum in the case of accidental death. The employers usually are able to transfer this risk by means of ordinary insurance. We also have in operation at the present time federal and state plans for providing workers with a measure of protection against the risks of unemployment and old age. The details of these various measures for the protection of the workers need not concern us in the study of economic principles, but it may be said that all three of the types of measures which have been mentioned provide rather small benefits for the workers, and that under the accident and unemployment protection the benefits continue only for a relatively short period of time. The unemployment and old-age measures afford protection to only a part of the workers of the country, since many groups of workers are excluded, and we have no general system of social insurance against the risk of occupational disease or illness. On the whole, therefore, the workers still bear a large part of their own risks, though our present protective devices may become very helpful if they are extended considerably.

SPECULATION

The Nature of Speculation in Commodities. We introduce some discussion of speculation at this point both because speculation may have a direct effect on some of the risks borne by enterprisers and because it is through speculation and speculative markets that enterprisers are able to hedge and thus protect themselves against certain risks of loss which might occur as the result of changing prices. Speculation in commodities consists of attempts by individuals to make money by buying and selling commodities whose future prices are uncertain. Since price changes are certain to occur for most commodities and since the direction and size of such changes can seldom be foreseen accurately, someone must bear a risk of loss in connection with these price movements. Businessmen can bear this risk, if necessary, but they are very willing, when it is possible, to allow professional risk bearers, or speculators, to carry the burden. Speculation has most significance in the case of commodities which are dealt in on organized exchanges, such as the Chicago Board of Trade, which deals in grains. Organized exchanges perform a number of functions: they furnish a place for

trading, regulate transactions, help in settling disputes, set up uniform grades for the commodities dealt in, and provide a system of inspection.

Exchanges usually permit the buying and selling of future contracts, or contracts calling for the delivery, at some definite future time, of a certain quantity of a certain grade of some commodity at a predetermined price. On these exchanges individual speculators are in a position to make money if their judgment of future price changes is more accurate than that of the market as a whole. Suppose, for example, that wheat is plentiful at a given time (say in February) but that coming crops of wheat are going to be unusually small. These coming short crops, foreseen by speculators and others, will tend to keep the present price of wheat higher than it would have been on the basis of present supplies alone, and the price in February of wheat for July delivery will be much lower than the price which would have prevailed in July in the absence of speculation and dealing in futures. Wheat might have been selling for 80 cents a bushel on the basis of present supplies alone, while the price might need to go up to $1.20 per bushel in July on the basis of supplies which would otherwise have been available at that time. On the grain exchange on February 1 we find that the actual price of present wheat of the given grade is, say, 97 cents per bushel, while the price at that same time for the same grade of wheat for July delivery is $1.03 per bushel, the 6 cent spread between these two prices representing roughly the cost of keeping a bushel of wheat from February to July.

Now if an individual speculator thinks that the February 1 price of $1.03 for July wheat is too low, and that the actual price of this wheat on July 1 will be higher, he will buy future contracts calling for the delivery of this wheat on July 1. If his judgment is correct, he will make money later on by selling his future contracts for July wheat, to someone who needs such wheat, at a price higher than $1.03 per bushel. The speculator, of course, does not want the wheat himself and is not prepared to take delivery of it. If, on the other hand, the speculator thinks that the February 1 price of $1.03 for July wheat, as determined by the actions of all buyers and sellers, is too high, and that the actual price on July 1 will be lower, he will sell July wheat "short." That is, he will sell future contracts calling for the delivery of wheat on July 1 at $1.03 per bushel and, if his judgment is accurate, he will make money later on by fulfilling these contracts at the lower price for July wheat then prevailing.

The Functions of Speculation in Commodities. Such activities of large numbers of speculators have certain rather valuable functions to perform in our economic system, provided that the speculative activities are carried on by well-informed, professional speculators and that their activities are honest and aboveboard. In the first place, speculation tends to minimize and smooth out fluctuations in the prices and consumption of commodities. In our preceding illustration, the price of wheat might have continued at about

80 cents per bushel during the winter and spring, in the absence of speculation, and consumption might have continued heavy on the basis of large present supplies. Then in the summer, when the shortage of wheat was apparent to all, the price would have skyrocketed to $1.20 per bushel and consumption would have been severely restricted.

In the face of coming shortage, the actions of speculators in buying wheat for future delivery tend to force the price up smoothly and gradually. The rising price, moreover, tends to hold down consumption at the present time so that the future shortage will not actually be so great as it would otherwise have been. Later on, when the short crops are in, the price will not have to be so high as it would have been in the absence of speculation and dealing in futures, nor will consumption be so restricted as it would otherwise have been. If, on the other hand, future supplies of wheat are going to be much larger than at present, so that the price will have to come down, the activities of speculators in selling wheat short will tend to depress the price smoothly and gradually, and the falling price will tend to encourage consumption. Later on the surplus will not be so great nor the price so low as would otherwise have been the case.

The actions of speculators also tend to stabilize prices between different markets for the commodities in which speculation occurs, so that the price of wheat in one market tends to differ from the price of the same grade of wheat in another market by not more than the cost per bushel of transporting the wheat between the two markets. If a greater price differential exists at a given time, it will tend to be reduced as individuals buy wheat in the market where its price is low and at the same time sell it in the market where its price is high. Finally, to the extent that speculation stabilizes prices, it lessens the risks of business enterprisers who must make use of the commodities in which speculation occurs. These enterprisers are bound to be exposed to some risk of loss through price changes when they use the commodities as raw materials or stock in trade, and anything which tends to minimize price changes will operate to reduce this risk.

When speculation is engaged in heavily by untrained, nonprofessional speculators, these advantages will not be completely realized. Uninformed outsiders who merely take a flier in the market are engaging in plain gambling. Betting on a turn in the price of wheat does not differ significantly from betting on the turn of a card. Such speculators are, of course, extremely likely to lose their money, and this result is especially likely if other operators in the market spread misleading rumors or misinformation. When nonprofessional speculators bet on the prices of commodities, their notions concerning future price changes have as much chance of being wrong as right. They may buy long when the price of a good needs to go down or sell short when it needs to go up. Such actions work against the tendency of professional speculation to have a stabilizing effect on prices.

Heavy speculation by outsiders at great cost to themselves, and sometimes to others, and the unethical activities of unscrupulous operators apparently make up the price which society must pay to obtain the benefits of legitimate speculative activity, for attempts to curb or eliminate these evils are likely to interfere significantly with the ordinary functioning of speculative markets.

Speculation in Securities. In our economic system there is also a great deal of speculation in the securities issued by corporations. Many changes in the situation of individual corporations and in general business conditions are able to affect the earning power of corporations and indirectly the prices of their securities. Thus the future prices of these securities are most uncertain and are subject to rapid and extensive changes, so that a fertile field for speculation exists. As in the case of speculation in commodities, speculators are able to make money by buying and selling corporate securities if their judgment of future price movements is superior to that of the market as a whole. Trading in securities takes place largely on organized markets or exchanges such as the New York Stock Exchange, the New York Curb Exchange, and other exchanges in principal cities.

The functions of speculation in securities are not exactly the same as those of speculation in commodities. For one thing, securities do not appear in crops at various seasons of the year, and speculation in them does not have the effect of smoothing out seasonal variations in supply and consumption. Again, since securities do not constitute the raw materials or stock in trade of ordinary productive establishments, any stabilization of security prices which results from speculation will not be likely to have the effect of lessening the risks of most business enterprisers. Finally, it is not usually possible for business enterprisers to hedge their regular productive operations by buying and selling corporate securities.

However, conditions on the security exchanges undoubtedly have some effect on the issuance of securities, and the operation of the exchanges furnishes a continuous market for listed securities—a market in which it is always possible to buy or sell—and this is of considerable value to investors. The existence of a continuous market contributes much to the easy transfer of the ownership of corporate securities without which the corporation would undoubtedly find it more difficult to obtain large quantities of capital funds from investors. The activities of speculators are important for the existence of a continuous market, for there are ordinarily speculators on both sides of the market—that is, some who are willing to buy and others who are willing to sell. Any price for a security on an exchange is a result of a compromise of the conflicting opinions of buyers and sellers on the market. That is, some operators think the existing price is too low while others think it is too high. If all operators considered the existing price of a security too high, the price would fall, and in the con-

verse case the price would rise. The behavior of security prices on the exchanges doubtless furnishes some guidance to investors in deciding which securities are worth purchasing and when they should be purchased.

Speculation in securities is also thought by some to have the effect of stabilizing security prices and minimizing fluctuations in these prices. Professional speculators are supposed to know or suspect when changes in general business conditions or in the condition of individual corporations will produce future changes in corporate earning power and security prices. If they think the price of a certain security will go up later on, they buy this security for holding, and their addition to the demand for the security causes its price to start upward at once. Later on, when the price would have gone up anyhow, the price change has been largely accomplished and little further readjustment is necessary. In the absence of speculation, the price might have remained at a low level for some time and then have gone up suddenly and sharply. In similar fashion, if speculators think that the price of a certain security will go down later on, they sell the security short and its price starts downward at once, instead of remaining high and then suddenly declining. Thus, speculation is said to have the effect of *discounting* future changes in security prices and of causing these changes to come about smoothly and evenly. While these effects on security prices may not mean much in connection with ordinary productive activities, they may be important to both businessmen and bankers in connection with credit operations. That is, loans are often made with stocks and bonds as security and the probability of sudden large changes in the prices of stocks and bonds might operate to discourage such loans.

On the other hand, the ease with which securities can be bought and sold on the exchanges and the widespread desire of people to make a little easy money have led at times to an immense volume of speculation in securities by poorly informed outsiders. This dabbling in the market by the general public has often interfered greatly with the normal functioning of the security exchanges. Unscrupulous operators have frequently spread rumors and misinformation in the hope of producing results on the exchanges which would be to their own advantage, and powerful operators have sometimes been able to manipulate security prices as desired through their own buying and selling activities. The so-called manipulative practices of market operators have not only caused great losses to outsiders who were taking a "flier" in the market but have produced fluctuations in security prices when none would have been necessary under ordinary conditions. As a result it is sometimes suggested that speculation in securities, on the whole, has little if any stabilizing effect on security prices. For more than a decade now our security exchanges have operated under federal laws and regulations designed to control or eliminate a number

of manipulative practices, but this aim is very difficult to accomplish without interfering with the normal functioning of the exchanges.

HEDGING

The Principle of Hedging. The existence of commodity exchanges, where dealing in futures is permitted, makes hedging possible. By means of hedging an enterpriser can protect himself to a considerable extent against one of his chief uninsurable risks—the risk of price changes affecting his commodity during the production period. The general notion of hedging is simple enough, for it involves, in a sense, betting both ways on an uncertain event. A student who bets on football games may hedge. If he places a bet on his own team at even money a week before a big game but, in the next two or three days, he learns some things about his own team or their opponents which make it seem likely that this bet will lose, he can hedge by placing a bet of the same size at even money on the opposing team. He can then view the game with equanimity, for he will not lose money regardless of the outcome.

A manufacturer who buys a certain quantity of a raw material to make a finished product which will be ready for sale some time in the future is also betting, in a sense, on price changes. If the price of his finished product tends to vary with the price of the raw material, he has placed himself in such a position that a rise in the price of the raw material and finished product will benefit him, while a fall in these prices will cause him a loss. Whether he so intends or not, he is betting that the prices in question will not go down. If he does not relish being in this risky position, the obvious remedy is to make another bet that these prices *will* go down. This he can do by selling short a quantity of the raw material equal to that which he is using for manufacturing purposes. Such a short sale protects the manufacturer to a considerable extent against the risk of price change. If the price of the raw material goes down and takes the price of the finished product with it, he loses money on the sale of his finished product but makes an equivalent profit on his short sale of the raw material. If the prices move upward, he loses money on his short sale but makes an extra profit on the sale of his finished product. Hedging thus involves a simultaneous purchase and sale of some raw material in such a way as to guard against the risk of loss through price change.

An Example of Hedging. Since the manufacturer tends to break even on his hedge, students are often puzzled to know how he ever makes any money. The following example, therefore, is intended to show both how a manufacturer may hedge and how he makes money even when he hedges. Let us consider the case of a flour miller and make certain arbitrary assumptions about his business. Suppose that his production period is one month and that he buys wheat on the first of one month intending to sell the re-

sulting flour on the first of the next month. Assume that it takes five bushels of wheat to make a barrel of flour, that the price of flour varies with the price of wheat, and that the miller is unable to sell his flour in advance of production.

It is the first of the month and the miller is considering the purchase of 5,000 bushels of wheat with which to manufacture 1,000 barrels of flour for sale next month. He knows the price of the desired grade of wheat to be $1.00 per bushel at present so that the cost of the material will be $5.00 per barrel of flour. He estimates his other expenses of production at $2.00 per barrel of flour on the basis of the projected volume of production. The kind of flour which he is to make now sells at $8.00 per barrel and, since his estimated costs will be $7.00 per barrel, he knows that, if the price of the flour will hold still for one month, he will have an income of $1.00 per barrel, or $1,000 altogether, for himself on the month's operations.

However, during the month he will be exposed to an important price risk if he does not hedge. If, by the first of next month, the price of the wheat has advanced to $1.25 per bushel and has taken the price of the flour to $9.25 per barrel, his 1,000 barrels of flour will sell for $9,250, which will leave an unexpectedly large margin of $2,250 above his assumed money costs of $7,000 for this quantity of flour. On the other hand, if the price of the wheat should drop to 75 cents per bushel during the month and take the price of flour to $6.75, his total receipts for 1,000 barrels of flour will be only $6,750. In this case he will not only lose his own expected income of $1,000 but will be out of pocket some $250 on his money expenses of production. The miller may or may not want to take this chance. Whether or not he hedges will depend upon his temperament, the scale of his operations, the recent trend in the price of wheat, and other matters.

Let us suppose that he decides to hedge. On May first he *buys* 5,000 bushels of the desired grade of wheat for milling. At the same time he *sells*, through some broker on the Chicago Board of Trade, a *future contract* calling for the delivery of the same quantity of the same kind of wheat on July first. His actual wheat is bought at $1.00 per bushel, while the future wheat is sold at whatever price is prevailing for July wheat on May first, say $1.02 per bushel. The purchaser of his future contract is some speculator who thinks that the May first price of $1.02 for July wheat is an underestimate of what the price will actually be in July. Clearly there must be some such speculators for, if many speculators thought the $1.02 price to be an overestimate and no one thought it to be an underestimate, the price of July wheat on May first would tend to be less than $1.02 per bushel.

Having concluded these two transactions, the miller is relatively safe. On the first of June he sells his 1,000 barrels of flour at whatever price is prevailing then and also proceeds to terminate his future contract for the delivery of July wheat. If the price of actual wheat has gone down to 75

cents per bushel during the month of May, he sells his flour for $6.75 per barrel, or $6,750 altogether, and thus loses his expected income of $1,000 as well as $250 of his expenses of production. However, if the price of actual or "spot" wheat has gone down, the price of July wheat will have gone down with it, since speculation keeps the two prices closely adjusted to each other. He can therefore get out of his future contract for July wheat at, say, 77 cents per bushel, whereas he sold the July wheat at $1.02 per bushel, and he makes a gain of 25 cents per bushel, or $1,250 altogether, on his 5,000-bushel short sale. This cancels his $250 loss on the flour and also gives him his expected income of $1,000 on the month's operation. On the other hand, if the price of wheat has gone up to $1.25 per bushel during the month, he makes a total income of $2,250, as previously shown, by selling his flour at $9.25 per barrel. However, he loses $1,250 on his short sale of July wheat, since he will have to get out of this contract at about $1.27 per bushel on 5,000 bushels, whereas he sold this contract at $1.02 per bushel. This loss of $1,250 on the short sale combined with a gain of $2,250 on the flour limits his net income for the month to the expected $1,000.

Thus the miller protects himself on his manufacturing operations by selling something which he does not have (July wheat) to someone who does not want it anyway, since the purchaser of the miller's future contract is some speculator who would be greatly surprised if actual wheat were ever delivered to him. As to how the miller terminates or gets out of his future contract on June first when he sells the flour, it is probable that he would simply pay 25 cents per bushel to the speculator or collect the same amount from him. In any case, however, the miller could get out of his short sale on June first by buying a future contract for 5,000 bushels of wheat deliverable on July first and turning this contract over to the speculator to satisfy the short sale, so that the miller cannot be kept from getting out of his short sale when the time has come to sell his flour. One should not feel too sorry for the speculator who bought the miller's future contract for July wheat, even when the price of wheat goes down and apparently causes the speculator a large loss on the transaction. The speculator, in buying the future contract, thinks that the price of wheat is going up and when, after a few days, he sees the price of wheat actually going down, he will also sell July wheat short in order to protect himself. The speculator who buys this new future contract will later sell short to protect himself too, and so on, so that the actual loss will probably be spread lightly over a number of speculators.

Other Considerations. This illustration of hedging is somewhat oversimplified, and some of the assumptions which were made concerning the milling business were clearly hypothetical, if not indeed contrary to fact, but the illustration shows definitely the principles involved in hedging, the methods of operating the hedge, and how the manufacturer who hedges can

still secure a net income for himself. However, too much significance should not be attributed to hedging as a means of protection against the risk involved in price changes, for it is not a device which can be used in all fields of economic activity. Hedging is for the most part limited to types of production using some raw material dealt in on an organized commodity exchange which permits dealing in future contracts. There are not many such materials. A manufacturer cannot hedge by selling short some raw material which he does not use in his operations, by selling securities short, or indeed by any method other than that of dealing in future contracts for his own raw material. Moreover, successful hedging presupposes a rather direct relationship between the price of the raw material and the price of the finished product, so that the two prices tend to change together. If the price of the raw material can change considerably without great effect on the price of the finished product, conditions are not favorable for hedging. The price of flour is supposed to vary rather directly with the price of wheat, but the price relationship between raw material and finished product is much less satisfactory for hedging purposes in other lines of production.

Even when hedging can be carried on successfully, it should not be supposed that the enterpriser is perfectly sure to receive his expected net income as a result. Hedging offers protection only against the risk of loss due to changing prices and not against all risks. Machinery can break down, workers can restrict output or go out on strike, and many other things can happen to bring losses even to manufacturers who hedge. Furthermore, hedging does not give complete protection even against the price risk. In our previous illustration of hedging we assumed that the price of July wheat in May would go up or down at the same time and to the same extent as the price of actual or spot wheat. These two prices do tend to go up and down at the same time but the "spread" between them (assumed to be two cents in our illustration) can change. Such changes in the spread during the production period furnish a small chance for profit or loss on the hedge itself and the expected net income of the manufacturer would be slightly augmented or reduced in such cases. However, hedging does ordinarily give protection against most of the risk of loss through price change.

Hedging by Grain Elevator Operators. The use of hedging as a protective device is not by any means limited to manufacturers. Enterprisers who operate grain elevators buy up large quantities of wheat and other grains and hold them for a certain period of time before final sales are accomplished. Since grain elevators sometimes store a million or so bushels of wheat, their operators would be exposed to a considerable risk of loss through falling wheat prices if they were unable to hedge. Their hedges are similar to that of the miller which has already been described. When the elevator operators buy wheat for storage, they sell future contracts to de-

liver wheat. Sometimes, but not very often, the wheat which they are holding is used to fulfill the future contracts. In other cases, the wheat which has been held is sold for cash and the future contracts to deliver wheat are fulfilled by buying other future contracts. By hedging, the elevator operators put themselves in such a position that a fall in the price of wheat will cause a loss on the stored wheat but an offsetting gain on the short sales, while a rise in the price of wheat will give an unexpectedly large net income from the storage operations but a loss on the future contracts to deliver wheat. In either case, the effect of hedging is virtually to make the price of wheat stand still for the elevator operator, whatever may actually happen to the price, and to assure the operator, other things being equal, of his expected net income from his storage operations.

Other Hedging by Manufacturers. Sometimes the position of the flour miller or other manufacturer is exactly the reverse of that described earlier in this chapter and he needs protection against a rise in the price of his raw material rather than against a fall in the price of his finished product. Suppose that a miller has succeeded in selling flour to be delivered in certain quantities at certain times in the future at predetermined prices. In such a case he is sure of the price which he will receive for his flour but he is not sure of the price which he will have to pay for wheat when it is needed at various times in the future. He therefore buys future contracts calling for the delivery of wheat at the times in the future when milling will have to be started for the various batches of flour. In some cases he may actually take delivery of the wheat called for by these future contracts and use it for milling. In other cases the wheat called for by the future contracts may not be exactly what he needs for milling, and he will buy spot wheat for milling and at the same time terminate his future contracts.

If the price of wheat goes up in the future, the miller stands to lose money on his flour which must be delivered at a fixed price, but he will make a gain on his long purchases of wheat for future delivery. If the price of wheat goes down, he makes a larger net income than he expected from his milling operations but sustains a loss on his long purchases of wheat for future delivery. In either case, the effect of a successful hedge is to assure the miller, other things being equal, of the net income which he would have made from his sales of flour had the price of wheat not changed at all during the life of his contracts for the delivery of flour.

QUESTIONS AND PROBLEMS

1. "While economic activity has always been risky, many of the risks to which business enterprises are exposed today have developed to their present importance only in comparatively recent times." Explain.

2. Why is it usually argued that landowners and capitalists are not very important as risk bearers in our economic system?

3. Do workers or business enterprisers bear the heavier risks in our economic system? Explain.

4. "Problems of risk and risk bearing are of little significance in non-capitalistic economic systems." Show whether you agree.

5. "Risks in economic activity must either be borne or be transferred to someone else, since it is impossible to reduce or eliminate them." Do you agree? Explain.

6. Is it possible to insure against a risk which affects only a few people? Why?

7. "Insurance results in the substitution of a small certain loss for a larger uncertain loss." Explain.

8. "Insurance permits uncertainty for the individual to be converted into relative certainty for a group of individuals and may result in a reduction of risk as well." Show whether you agree.

9. What are the problems involved in insuring against sickness, unemployment, and loss of bank deposits?

10. Explain the general nature of speculation in commodities.

11. Why is organized speculation in commodities sometimes considered to be socially beneficial? Explain.

12. "The functions performed by organized speculation in securities are not the same as those which result from speculation in commodities." Explain.

13. What disadvantages result from speculation in commodities and securities? Do they offset the advantages? Explain.

14. Give a clear and precise account of a hedging transaction.

15. "A hedge may involve either the buying or the selling of future contracts." Show whether you agree.

16. What is the specific purpose of hedging? That is, why is a manufacturer better off if he hedges than if he does not hedge?

17. "Hedging guarantees that a manufacturer will make a profit from the production and sale of his goods." Do you agree? Explain.

18. "The manufacturer who hedges may make money even though he breaks even on the hedge itself." Explain.

19. "Since hedging affords considerable protection for the manufacturer, it is peculiar that all manufacturing enterprises do not engage in hedging transactions." Discuss.

20. It is easy to see why a manufacturer may want to hedge, but how can he be sure of finding a speculator to buy or sell the necessary future contract? Explain.

REFERENCES FOR FURTHER READING

Boyle, J. E., *Speculation and the Chicago Board of Trade.* New York: The Macmillan Company, 1920, Chaps. 4–5.

Converse, P. D., and Huegy, H. W., *Elements of Marketing.* New York: Prentice-Hall, Inc., 2nd ed., 1940, Chap. 12.

Erdman, H. E., *The American Produce Markets.* Boston: D. C. Heath and Company, 1928, Chaps. 9–10.

Hardy, C. O., *Risk and Risk-Bearing.* Chicago: University of Chicago Press, 1923, Chaps. 1–5, 11–12, 17–18.

Magee, J. H., *General Insurance.* Chicago: Richard D. Irwin, Inc., 1936, Chaps. 1–7.

Marshall, Alfred, *Industry and Trade.* London: Macmillan & Company, Ltd., 1920, Book II, Chap. 5.

X

Introduction to Price Determination

Direct and Indirect Exchange. The modern process of production, as described in several preceding chapters, is roundabout, specialized, and large scale, and a tremendous volume of exchanges is required to get economic goods into the hands of final consumers. The exchanges of products which occur in an economic system may be either direct or indirect in character. Direct exchange, which is also called barter, involves the swapping of economic goods (commodities or services) directly for each other. In indirect exchange, the individual gives up commodities or services in return for a quantity of some medium of exchange (money or credit) and then uses this purchasing power to acquire the other commodities and services which he desires. This method is called indirect because an additional step is introduced into the process of exchange and an intermediary comes between the economic goods which are to be traded for each other. Direct exchange is frequently used under primitive economic conditions and, it must be admitted, is much better than no exchange at all. However, as an economic system develops and progresses, people usually find something, or perhaps several things, which may be used as a medium of exchange, and direct exchange gives way to indirect. The reasons for this change are found in the disadvantages of barter.

Obstacles to Direct Exchange. In the first place, direct exchange is difficult to carry on because individual buyers and sellers may not desire each other's products. If a man has some tomatoes to dispose of and desires a pair of shoes in return, he must find someone who desires tomatoes and is willing to give shoes in exchange. It might be easy to find someone who would accept tomatoes or someone who would part with shoes, but difficult to discover both in the same person. Then, too, there is the problem of divisibility under barter, for many things cannot be divided without impairing or destroying their usefulness. If an individual wishes to sell a cow and desires in exchange at the time only something which is worth the sixteenth part of a cow, what is he to do? He might, of course, dispose of the animal and accept in return the article of lesser value plus some prom-

ises on the part of the cow-buyer to give him some other articles at a later time, but this would really be to introduce a form of credit into the transaction. The same problem would be seen if a lawyer needed a pair of shoe laces under direct exchange and had only his legal services to offer to the seller. The owner of the shoe laces might well scoff at the idea of parting with them in return for a minute of legal advice from the lawyer. He would point out that such a small portion of the lawyer's services would be valueless to him, and yet the lawyer could scarcely afford to offer more.

The problem of expressing values is troublesome under direct exchange. Lacking any medium of exchange in terms of which the values of other things might be expressed, the seller of an article would have to be prepared to state the value of his good in terms of all articles which he would be willing to accept in exchange. Thus a tailor might have to decide that a suit of clothes which he had for sale would exchange for fifty bushels of potatoes, five hats, one hundred haircuts, half a radio, and so on. Finally, under barter, there is the problem of storing up values. If a man has a crop of ripe strawberries which must be sold at once but does not wish any other commodities or services in exchange at present, he is in a difficult position. The berries will be worthless if he keeps them until he wants other things, and if he sells them now he must take some other things which he does not really want at present. If he swaps the strawberries for some less perishable article and keeps the latter to trade later for some desired goods, he is really using the durable article as money and, if he exchanges the strawberries for a promise of other goods in the future, he is really falling back on the use of credit.

Advantages of Indirect Exchange. These various difficulties disappear readily when money or credit is used as a medium of exchange. The tomato seller has merely to find someone who will give him money or credit for his good and then find some other party who will take the money or credit in exchange for shoes. This process is usually much easier to carry out than that of finding a single person who will accept tomatoes and give shoes. The cow or the legal services may be sold for money, which is readily divisible into small units for the purchasing of shoe laces, ice-cream cones, or other articles of small value. With money serving as a sort of common denominator for expressing the values of economic goods, the tailor may price his suit at fifty dollars and everyone will know what that means and can determine the value of what he has to sell in relation to the suit. Finally, perishable articles such as strawberries may be sold for money and the money stored until a later time when it will be used to procure other economic goods in exchange. To be sure, when the money is finally spent it may not have exactly the same power to command other goods in exchange which it had when it was received, but money nevertheless furnishes a relatively satisfactory device for storing values through time.

THE GENERAL SIGNIFICANCE OF PRICES

A system of indirect exchange based on money, credit, and prices operates much more successfully than one of direct exchange or barter, but it would be a mistake to think of prices as being important solely in connection with facilitating the exchange of goods. In a capitalistic economic system such as that of the United States, prices and price relationships constitute the basis on which many very important economic decisions are made.[1]

The Control of Production. In the first place, price relationships control the kinds and quantities of economic goods which are produced in a capitalistic system. Prospective business enterprisers, in trying to decide which commodities and services are worth producing and in what quantities, are attracted to the production of economic goods whose prices stand in favorable relationship to their costs of production. Existing enterprises expand their rates of production when the prices of their products rise above the level of cost of production, and slow down production or even cease producing temporarily when the prices of their products fall below the level of cost of production. Firms are induced to leave industries when the prices of products remain unfavorable over considerable periods of time and to take up their abode in other industries in which more favorable cost-price relationships exist. Thus increases and decreases in production, expansions and contractions of industrial capacity, and conversions and reconversions of productive facilities normally occur on the basis of the prices of finished goods and productive agents in our economy.

The Allocation of the Agents of Production. When individuals have land or capital goods for sale, they ordinarily sell these productive agents to the individuals or firms which will pay the largest prices for them. Similarly, when individuals have land or capital for rent or savings to invest, they ordinarily turn these agents of production or savings over to the enterprises which will undertake to pay the largest sums for their use, after due allowance for varying risks is made. Even in disposing of their own services as workers, individuals are greatly influenced by the prices (wages) which will be paid for these services in the various occupations for which they can qualify, although other considerations may also be of some importance. These activities of individuals, in selling their productive agents or the uses of these agents, serve to allocate the existing limited supplies of the agents of production among industries and businesses and to coordinate the uses of the agents with the desires of individual consumers as expressed through prices in the market. The latter result is achieved because the ability of enterprises to pay high prices for productive agents is related to their ability to get high prices for their final products.

[1] The significance of prices in general under capitalism and the functions which are performed on the basis of price relationships have been discussed in some detail in Chapter IV.

The Apportionment of Finished Commodities and Services. Since consumers' goods and services never exist in quantities which are adequate to satisfy fully the wants of all consumers, some method of apportioning limited quantities of finished economic goods among those individuals who want them is always necessary. This task is usually performed by increases and decreases in the prices of these goods. An increased price for an economic good will reduce the number of persons who will demand it at any given time or the quantities which they will purchase, while a lowered price will have opposite effects. No matter how small or large the available quantity of an economic good may be, the effective market demand can always be increased or decreased to a corresponding extent through a rise or decline in the price of the good. And if the rises and declines in prices which are necessary for this process are adequate to force prices above or below the level of cost of production, enterprisers will be led to take appropriate action with respect to the volume of production in longer periods of time.

Controlling Total Amounts of Productive Agents. Finally, we must note that prices or price relationships control, where possible, the total amounts of agents of production in existence in the long run. High wages in a particular occupation are likely to induce more individuals to prepare themselves for that field of endeavor, instead of other occupations for which they might qualify, even though no effect of high wages on total population or labor supply may be forthcoming. High rates of earnings for particular kinds of capital goods will certainly lead to the widespread duplication of these goods under competitive conditions. A high rate of interest for savings, after due allowance for risk, is likely to induce many people to forgo the immediate consumption of portions of their money incomes in order that their savings may be increased. In converse fashion, low rates of wages may be expected to discourage workers from seeking or preparing for a certain occupation, low rates of earnings will induce enterprisers to fail to replace existing capital goods when they wear out, and a low net rate of interest is likely to discourage saving, in the long run.

Thus we see that prices play a vital part in the functioning of our economic system and in the whole economic process of satisfying human wants through the use of scarce resources of production. Any change in the price of a finished economic good or a productive agent may have far-reaching consequences. If the price of a consumers' good increases, the consumption of the good is likely to be decreased and the substitution of other goods for it may be encouraged. Furthermore, if the price remains at a high level, production of the good may be stimulated, additional productive facilities may be diverted to the industry, and there may even be an increase in the total quantity of productive agents suitable for producing the good. Similarly, if the price of a certain productive agent, or one of its grades, increases, cost-price relationships in industries and businesses which use the

agent may be affected, the production of various finished products may be reduced, the quantities of the agent which enterprisers will demand may decline, the substitution of other productive agents for the one in question may be encouraged, and in time an increased quantity of the high-priced agent may be produced and made available. In view of the effects which the changing prices of finished economic goods and productive agents may have, it is clear that the factors or forces which determine these prices are of great importance. It is these factors which we are to study in the present chapter and in several following ones.

VALUE AND PRICE

The present section of the study of the principles of economics is called both the Theory of Price Determination and the Theory of Value. The fact that these terms are used interchangeably suggests that a fairly close relationship must exist between values and prices, and this is indeed the case. *The value of an economic good is its power to command other economic goods in exchange.* Thus, as we noted in connection with barter, the value of a man's suit of clothes might be fifty bushels of potatoes, five hats, one hundred haircuts, or half a radio, if such were the quantities of these economic goods which the suit of clothes could command in exchange for itself. *The price of an economic good is its power to command money in exchange.* In the value situation just described, if the price of the suit of clothes is fifty dollars, the potatoes, hats, haircuts, and radios must have prices of one dollar, ten dollars, fifty cents, and one hundred dollars per unit, respectively. Under barter exchange, economic goods have values but not prices. Under indirect exchange, they have prices and also (indirectly) values in terms of other economic goods.

The Relations of Value and Price. At any given point in time, price and value have almost the same significance for, if an economic good will command a certain amount of money at a given time, it will also command equivalent amounts of other economic goods at that time. However, over a period of time, changes in the prices of individual economic goods may not measure at all accurately the changes which occur in the values of these goods, since prices are stated in terms of money and the value of money itself changes frequently. If sirloin steak sold at retail for 40 cents per pound in 1929 and for 32 cents in 1932, its price decreased by 20 per cent but its power to command other goods in exchange actually increased during this period, since retail prices in general in 1932 were more than 30 per cent lower than in 1929. Or suppose that a shirt of a certain quality sold for $2.50 at the beginning of 1941 and for $2.75 at the end of 1944. The price of the article increased by 10 per cent during this period, but its value, or power to command other goods in exchange, actually decreased because prices in general increased by about 25 per cent during the same period.

The price of a given economic good may even fall when prices in general are rising or rise when prices in general are falling. Clearly, then, in a period of time in which the general price level and value of money are changing, the change which takes place in the price of an economic good is very likely to represent inaccurately the extent of the change which occurs in its value, and may not even indicate accurately the direction of the change in its value.

How, then, can we use value and price to mean virtually the same thing for our present purposes? The answer is that in this entire section of the study of economics we assume that the general price level and the purchasing power of money remain constant. Under this assumption, of course, changes which occur in the prices of economic goods represent equivalent changes in their values, and it matters little which term is used. The making of this assumption does not mean that we think that the purchasing power of money and the general price level actually remain constant in practice in our economic system, for we know that these things change frequently and to a considerable extent. Their changes are of great importance to numerous classes of people and they are quite worthy of our study, but we must defer this study to a later chapter. What we are doing here is merely to eliminate temporarily from our consideration changes in the purchasing power of money and the general price level in order that we may concentrate attention on the other factors which influence the prices of economic goods, for an attempt to consider at one time all the influences which affect the prices of economic goods would result only in confusion. This method of temporarily eliminating from consideration certain factors while other factors are studied is frequently used in other sciences as well as in economics.

In practice, since the purchasing power of money and the general price level do change, it is possible for all economic goods, or economic goods in general, to rise or fall in price. That is, all economic goods may come to command more or less money per unit than formerly. However, it is not possible for a general rise or fall in values to occur. Since values are expressed as ratios between economic goods, it follows that, if one economic good comes to command more of other economic goods in exchange, the other economic goods must command less of the first economic good in exchange than formerly. If suits of clothes come to command more pairs of shoes than formerly in exchange, then shoes command fewer suits of clothes in exchange than formerly. Increases in the values of some things involve decreases in the values of other things.

Having temporarily eliminated changes in the purchasing power of money and the general price level as factors affecting the prices of economic goods, we are in a position to give strict attention to the other factors which influence these prices. These other factors are the forces of de-

mand and supply. Under the assumption which we have made, the price of any economic good is determined by conditions of demand and supply, and changes in its price result only from changes in conditions of demand and supply. Of course, a great number of factors may affect the price of an economic good. Floods and droughts, wars and rumors of war, breakdowns of machinery, restrictive policies of labor unions, the invention of new commodities, transportation tie-ups, improved methods of production, and changes in the tastes of consumers are a few of the many factors that might affect the price of a certain economic good, but it should be noted that these factors can operate on price only through their effect on demand or supply, or both. However, we do not get much nearer to our goal of understanding price determination by learning that prices are determined by conditions of demand and supply, for we must spend several chapters in learning what demand and supply are and how they operate in different periods of time and under various conditions of the market. Before turning to an analysis of demand, we shall examine in some detail the various conditions which may prevail in the markets for economic goods.

CONDITIONS OF THE MARKET

The Nature of a Market. The term "market" often conjures up in the minds of students a definite market place, such as a farmers' produce market in the city, a stock exchange, or a supermarket for groceries and meats. These things are markets, of course, but the term as used in the theory of value is somewhat broader and less definite in significance. Thus a market may be defined as *any region or area, regardless of size, in which buyers and sellers are able to deal readily in a given economic good.* Some things have a market which is practically world wide in extent (such as the precious metals, securities, and staple commodities like wheat, cotton, or wool). Such wide markets for certain economic goods depend upon a number of factors. The articles in question must be able to stand transportation. That is, they must ordinarily have a value which is rather high in relation to their bulk, and be relatively immune to deterioration while being transported. They must usually be highly standardized products, or products which can be dealt in satisfactorily on the basis of grading and sampling, and, of course, the demand for them must also be widespread.

For many years articles such as fresh fruits and vegetables were sold almost entirely in local markets because of their perishability. Now, however, methods of packing and transporting these goods have been greatly improved, and they can be bought and sold over comparatively wide areas. Articles which have a small value in relation to their bulk, such as cement or bricks, are sold for the most part in market areas of fairly limited size. The fact that the prices of doctors' and lawyers' services, and those of motion-picture performances, often vary considerably from one community

to another indicates that nonmaterial economic goods are bought and sold to a large extent in local markets.

The problem with which we are dealing at present, however, has to do not with the size of the market but with the conditions which prevail in the market. The principles involved in price determination differ to some extent according to whether the conditions which prevail in the market for a given economic good are those of competition, monopoly, or something in between these extremes. Here we are concerned primarily with supply conditions. While conditions of demand, in strict theory, could approximate those of monopoly or semimonopoly, it is usually safe to assume that the demand for an economic good is competitive in character. Certainly the most important cases of departure from competitive conditions are found in practice very largely on the side of supply.

Competitive Market Conditions. Although competition, monopoly, and other market conditions were discussed to some extent in Chapter III, it is necessary to enlarge on the earlier discussion at this point because of the great importance of these market conditions in connection with the theory of value. The first requirement for the existence of a competitive market is that there must be a large number of buyers and sellers of exactly the same economic good in the market. A "large number" here means so many buyers and sellers that any one buyer or seller can withdraw from the market, or a new buyer or seller can enter the market, without affecting the price of the good in question. It also means that any one buyer can purchase as much or as little of the good as he pleases, and any one seller can sell as much or as little as he desires, without affecting the price of the good. The sellers must be selling exactly the same economic good in order that there may be no reason for buyers to prefer the product of any particular seller to those of other sellers. As soon as there is some reason for buyers to prefer one seller's product over another's, the individual sellers begin to have some trace of monopolistic power.

Another requirement for a competitive market has to do with the independence of the individual buyers and sellers. In such a market each individual buyer or seller is supposed to reach a decision as to the quantity of the economic good which it will be worth while to buy or sell, on the basis of the actual or prospective price of the good, without cooperating with other buyers or sellers. There must be no agreements, combinations, or conspiracies among buyers for the purpose of restricting purchases and bringing the price down or among sellers for the purpose of restricting production or sales and forcing the price up or keeping it high. Finally, in a competitive market, both buyers and sellers must have reasonable knowledge and information as to market conditions governing the particular economic good. This does not mean that every buyer must qualify as a technical expert in connection with each good he purchases or that each seller must know all

there is to know about the production and sale of his good. It means only that the buyers of an economic good should be sufficiently wide-awake so that they will not pay $10.00 or $7.50 per unit for an economic good which can be purchased for $5.00, and that sellers should know enough about their sales opportunities so that they will not take 50 cents per unit for an economic good for which the prevailing price is 75 cents or $1.00 per unit.

Strictly competitive conditions, as we have defined them, do not exist very commonly in practice in our economic system. However, conditions in the market for a number of agricultural products approximate the competitive ideal quite closely. That is, there are hundreds of thousands or millions of growers of these products, very numerous dealers or middlemen, and, of course, millions of final purchasers or consumers. The individual buyers are quite independent of each other; and so are the growers except when the government steps in with some plan for the control of agricultural production and prices. Most of the products have several grades or types, but one farmer's product of a given grade is scarcely distinguishable from another's. There is very little advertising of most of the products, and entrance to or exit from any one branch of agricultural production is relatively easy for the individual enterpriser. In addition to agriculture there may be a few other fields in which markets are more nearly competitive than anything else.

Monopolistic Market Conditions. Conditions of competition constitute the limiting case at one end of the scale of market conditions. At the other end are conditions of monopoly. In value theory conditions of monopoly in a given market refer to a situation in which the supply of an economic good is controlled by a single seller, or a group of sellers acting as one, while competitive conditions continue to prevail on the demand side of the market. The aim of the monopolist in operating his business is not essentially different from that of any other enterpriser, since he merely tries to derive as much net income from his business as possible, but since he does not have to worry about the actions of other sellers of the same product, he is in a better position to accomplish this aim than is an enterpriser in a competitive industry. The monopolist does not control the demand for his product and he must therefore adapt his production and sales to this demand. He can set his production and sales at any figure within his capacity, but the price per unit which he can obtain will then depend upon demand. On the other hand, he can set his price per unit at almost any level he chooses, but the amount of the good which he can sell will then depend upon demand.

Even if his knowledge of demand conditions is very good, and he has a better chance than numerous competing sellers to achieve such knowledge, because of certain checks on his power the monopolist may not be perfectly free in practice to charge the price which will maximize his net income from the business. If the monopolist charges a very high price per unit for

his good, substitute products may be developed and used. Even though a single seller, or a group of sellers acting as one, has complete control over the supply of a particular economic good, a monopoly can scarcely be said to exist if there are reasonably close substitutes for this product. Again, while the monopolist may have complete control over the production and sale of an economic good at a given time, he may find that, if the profits which he makes are extremely large, other firms will be attracted to exactly the same line of production and he will have competition on his hands. Finally, if the monopolist charges a very high price and the consumers of his product suffer greatly from his exercise of monopoly power, the government may step in and try to break up the monopoly or regulate the business which the monopolist carries on.

These checks on monopoly power may not all be effective in any given case of monopoly. While some monopolists need to worry about the possible development of substitute products, it is obvious that this consideration would not greatly concern a firm which possessed a salt, sugar, or tobacco monopoly. Again, a monopoly which owns or controls all domestic sources of raw materials for the production of a certain economic good, and is protected from foreign competition by the tariff, does not need to worry about the entrance of new firms into the industry even though its monopoly profits are very large. Even the danger of governmental interference may be discounted to some extent since the monopolist may be able to conceal his profits rather successfully and since the efforts of government to break up or control monopolies in the past have never been particularly effective.

The Extent of Monopoly. Cases in which a single seller, or a group of sellers acting as one, controls the supply of an economic good completely are comparatively rare in practice. However, we were told recently that "one company in each field controls all, or nearly all, of the nation's supply of aluminum, nickel, molybdenum, magnesium, shoe machinery, glass container machinery, and scientific precision glass, provides nearly all of the domestic telephone service . . . and operates all of the sleeping and parlor cars." [2] And, of course, there are hundreds of other cases in which a single firm or combination of firms dominates the production of individual economic goods and exercises a considerable degree of monopoly power even though it controls only 50 to 75 per cent of the available productive facilities.

Public utility companies supplying such economic goods as water, gas, or electricity often operate as complete monopolies in given market regions, but in these cases the companies operate as monopolies with the specific permission of the government and under regulation. These public utility companies are often called "natural monopolies." They are in fields of pro-

[2] Temporary National Economic Committee, Monograph No. 21, *Competition and Monopoly in American Industry*. Washington, D. C.: Government Printing Office, 1940, p. 69.

duction which require a heavy investment in fixed plant and equipment and in which competition would lead to a great duplication of productive facilities, heavy overhead costs per unit of product sold, and high rates to the consumers. Even if they begin to operate under competitive conditions they tend to reach the monopoly form in the end, for in an effort to attract business the various competing companies tend to cut rates on each other until all companies save one are eliminated or until the competing firms see the error of their ways and decide to combine. Under such conditions, it may be deemed wise to permit the industries to operate in the monopoly form although their activities are regulated by public authority. We should also note in passing that the single seller of electricity in a given region may have to compete with sellers of coal and gas, for example, in connection with the cooking and house-heating activities of consumers even though there is no close substitute for electricity for purposes of lighting.

Some economists speak also of monopolies which exist by virtue of patents and copyrights granted by the government. New inventions, new products, and new features of products may be covered by patents which confer on the holder of the patent, or those who will pay royalties to him, the exclusive right to make and sell the good in question. In some cases patents have been the basis for effective monopoly. However, it is likely that other firms will try to turn out products which duplicate the successful patented product as closely as they may without violating the law, so that the firm which owns the original patent is likely to find itself in a situation of monopolistic competition rather than one of monopoly. When General Motors patented a certain system of "no-draft ventilation" for automobiles, other manufacturers soon came out with other ventilating systems which resembled, but did not actually duplicate, the patented feature of the cars produced by General Motors.

In similar fashion, copyrights grant to individuals or firms the exclusive right to reproduce musical compositions, plays, and books or to use trademarks and advertising slogans in business. The copyright covering a book which becomes a "best seller," such as *A Tree Grows in Brooklyn* or *Arch of Triumph,* a popular song such as *Don't Fence Me In* or *The Atchison, Topeka, and Santa Fé,* or a play or other stage production such as *The Voice of the Turtle* or *Oklahoma,* may be highly valuable and a great source of income to its owner. Much the same thing may be said of trademarks and trade names, such as *Body by Fisher, It's Toasted, No Metal Can Touch You, Nabisco, Kodak,* or *Phillies.* However, while the owners of copyrights are much better off than they would be if other persons were allowed to reproduce and sell freely these songs, books, and plays, or use the same trade-marks and trade names, it is usually true that other sellers offer for sale roughly similar songs, books, and plays or use similar trade-

marks and trade names. Here again the situation which results is likely to be one of monopolistic competition rather than one of pure monopoly.

Duopoly and Oligopoly. Several conditions of the market lie in between the extremes of competition and monopoly. In one situation, known as duopoly, there are only two sellers of a given economic good while conditions on the demand side of the market remain competitive as usual. Of course, if the two sellers of an economic good act together to determine production, sales, and prices, the situation is one of monopoly rather than duopoly. Even when the two sellers are as independent of each other as possible, any change in production, sales, and price by one duopolist is of tremendous concern to the other, and each must always pay close attention to what the other is doing. Cases of duopoly are not very numerous in our economic system, but it was reported that, in 1937, pairs of firms controlled all or nearly all of the supply of certain economic goods in the domestic telegraph service, the importation of bananas, the production of plate glass, the production of glass bulbs, glass tubing, and rod for electric lamps, the production of electric accounting machines, the manufacture of railroad air brakes, the production of compressed oxygen and acetylene, and sulphur production.[3] However, in some of these cases, the pairs of firms may have acted together to create a situation of monopoly rather than one of duopoly.

In the market situation known as oligopoly, there are only a few sellers of a given economic good while conditions of demand remain competitive as usual. The few sellers of an economic good may act together to determine production, sales, and price, but in this case the situation is one of monopoly rather than oligopoly. Even in the case of genuine oligopoly the actions of any one seller with regard to production, sales, and price are of great significance to the other sellers, for any one firm is large enough to have an important effect on these matters from the point of view of the whole industry. Cases which are at least potentially oligopolistic are quite common in our economic system. In the case of 37 products, four firms accounted for the entire supply in 1937, and four firms turned out over 90 per cent of the supply in 164 cases. Altogether it was estimated that from two fifths to one half of the 1,807 products studied were produced in fields in which four firms controlled 75 per cent or more of the supply.[4] Not all of these cases were necessarily genuine oligopolies, however. In the case of some products, the few sellers may have acted together as a monopoly. In other cases, the few sellers may have produced differentiated products rather than exactly the same economic good. Where this was true, the situation was really more nearly one of monopolistic competition than of pure oligopoly.

[3] *Ibid.*, pp. 98–110.
[4] *Ibid.*, pp. 113–118.

Monopolistic Competition. Many of the market situations which lie in between competition and monopoly fall under the heading of monopolistic competition. This condition of the market is similar to competition in some respects. In the first place, the number of sellers (and buyers) in the market under monopolistic competition may be just as great as under competition. In some actual cases the number of sellers may be restricted, but this is not at all necessary to the existence of monopolistic competition. In the second place, monopolistic competition does not involve any collusion or agreements among the producers in an industry for the purpose of controlling production, sales, and prices. The actions of one monopolistic competitor may have more important effects on the other firms than the actions of one producer under competition have on other producers, but this results from the market situation itself and not from any cooperative activities on the part of the monopolistic competitors. Finally, the competitive condition of reasonable knowledge and information on the part of individual sellers is also present under monopolistic competition. In fact, the individual monopolistic competitor may be more likely to have adequate information concerning conditions of demand and the activities of other producers than the individual competitor.

Product Differentiation. The one necessary and decisive factor in distinguishing between competition and monopolistic competition is found in the fact that, in an industry operating under monopolistic competition, the individual firms and companies do not all produce and sell precisely the same economic good. By means of patented mechanical features, as in the case of automobiles or electric refrigerators, or through peculiarities in style, color, quality, or design, the individual sellers attempt to set their products apart from all others of the same general type and to attract and hold a large number of customers for themselves. One electric refrigerator will freeze only 64 ice cubes at one time while another will turn out 128, or some other number. One has a patented device for removing the ice cubes from the trays and molds, while another lacks this device. One has shelves on the inside of the door while the doors of others contain nothing more than instructions for operating. One has its motor underneath the body of the refrigerator, while another has its motor on top. Special containers for meat, vegetables, oranges, and eggs vary from one refrigerator to another. Where significant differences exist between the products of individual sellers in a general line of production, it is clear that ordinary conditions of competition do not exist.

But what are significant differences? Almost anyone might be willing to accept the physical features noted above as significant differences, but it is often stated that such differences also exist when substantially identical products are sold by various sellers under different trade names or trademarks, or in different packages and containers. Indeed, some writers sug-

gest that competition exists only when the locations of the sellers are equally convenient to all buyers, and when, from the point of view of the buyers, all sellers have equally desirable reputations or equally pleasing personalities. According to this same point of view, product differentiation exists if one store is air-conditioned and another is not, and the same thing is true if one restaurant employs waiters and another waitresses, or if one restaurant has pretty waitresses while another does not.

It is difficult to say just where product differentiation begins and ends, but, if all the factors just mentioned are accepted as product differentiation which results in monopolistic competition, it may seem that we have defined monopolistic competition so broadly that practically all markets in which perfect monopolies do not exist would have to be characterized by monopolistic competition. That is, it would be virtually impossible for genuine competitive conditions to exist in practice. So we must ask ourselves, for example, whether table salt in a blue container is significantly different from table salt with precisely the same physical characteristics which is sold in a red container. It is tempting to say that the answer depends upon the nature of the demand for the product and that, if the demand is really competitive, these two table salts constitute a single, identical economic good. One of the conditions of competition from the point of view of demand is that buyers should have reasonable knowledge and information concerning an economic good which they are buying. If this condition prevailed in practice, it would be considered very doubtful whether sellers could convince buyers that table salt in a blue container is significantly different from table salt of the same quality sold in a red container, and conditions of monopolistic competition would be limited to cases in which there existed some physical differences in the actual products of individual sellers in a given industry.

However, this attitude may represent a considerable oversimplification of the issue. Some people, for example, are much more sensitive to color than others. A man buying table salt of given quality may not care at all whether the container is blue or red, but his wife may greatly prefer a blue container to a red one, since she repeatedly sees it standing on the shelf at home. A man buying groceries may care little whether dried beef comes in a cellophane wrapper or in a glass jar, but his wife may greatly prefer the latter alternative since the jar will turn into an excellent jelly glass when the meat has been consumed. Air conditioning in a store may mean little to some individuals but much to others. To many people differences of the kinds just discussed may be just as important as the difference between an electric refrigerator which has shelves on the back of the door and one which does not.

The Individual Seller under Monopolistic Competition. The individual seller under monopolistic competition is in a distinctive position with regard

to the market for his economic good. His objective in operating his enterprise will not differ greatly from that of any other enterpriser since he, too, will be trying to maximize the net income derived from the business, but his position with respect to the possibilities of achieving this objective is different from that of the monopolist or competitive seller. Unlike the monopolist, he cannot proceed blithely to carry on his business without having to worry about any other sellers of his general type of economic good. On the other hand, he is not able, as is the individual seller under competition, to sell any quantity of his good within reason at an established market price, for he is in a field where product differentiation exists. The amount of his economic good which he can sell will depend upon the total demand for that general variety of good, on the specific variety of the product which he decides to sell, on the price per unit he decides to charge, on the specific varieties of the product which other sellers elect to produce, on the prices per unit which other sellers charge, and also probably on the outlay which he is willing to make for advertising his particular variety of the product in relation to the amounts other firms spend for the same purpose.

The Extent of Monopolistic Competition. If we accept patented features, other physical differences, distinctive packages and containers, and differences in trade-marks, trade names, and slogans as typical of product differentiation, it must be evident that very large numbers of economic goods are sold under conditions of monopolistic competition in our economic system. For example, although the clothes and shoes which the college professor puts on when he gets up in the morning may conform to current styles and designs, they have undoubtedly been sold under distinctive trade-marks, trade names, and slogans. The soap with which he washes and the powder or paste with which he cleans his teeth are produced according to patented formulas, sold in distinctive packages or containers, and protected by trade-marks and trade names which have been copyrighted. The breakfast food on his table, which in physical composition is not very different from other cereals, will certainly have been sold under a trade-mark and trade name, and in a distinctive package or container, with perhaps a patented inner container to keep the cereal fresh. Many other foods, though not all, are sold under some of these devices of product differentiation. The electric refrigerator in our professor's kitchen will have a variety of patented features, as previously reported, along with the usual trade-mark and trade name. The furniture which he uses at home, the automobile in which he rides to school, the desk at which he sits at his office, the fountain pen with which he writes, the watch with which he times his classes, and a hundred and one other things, are sold under distinctive trade-marks and trade names and are likely to have patented features and other physical differences. Thus the field of monopolistic com-

petition is very broad, although in some cases conditions of oligopoly cut across those of monopolistic competition.

The Approach to the Theory of Value. While the theory of price determination under conditions of monopolistic competition is not altogether new by any means, there has been in the last fifteen years a considerable burst of activity on the part of writers in this field which has served to bring the matter forcibly to the attention of economists. With regard to the presentation of the theory of value, there are now two opposing schools of thought as to the place of the theory of monopolistic competition. Some writers, greatly impressed with recent developments in the field of monopolistic competition, have presented the theory of value almost entirely in terms of this condition of the market. Having defined the term so broadly that it includes practically all market conditions save those of monopoly or duopoly, their presentation of the theory of value leaves out much of the traditional material with regard to price determination under competitive conditions because they think that this material is of scarcely any value to describe or even approximate what goes on in actual economic life.

Now it may be admitted that competitive and monopolistic conditions of the market, as we have defined them, are not often found in practice and that they are extreme cases. It is definitely a mistake today to present a theory of value which assumes that practically all markets for economic goods are actually competitive in character, with monopolistic market conditions existing in a few scattered and exceptional cases, and with nothing existing in between these two types of markets. On the other hand, it does not seem wise to concentrate attention altogether on the conditions which fall between competition and monopoly, for the description of price determination under monopolistic competition is too complicated to serve effectively as an introduction to the theory of value. It seems likely that the student will be better able to understand price determination under the conditions which fall in between competition and monopoly if he studies first the way in which prices are determined under these admittedly extreme and limiting cases, and it might still be possible to reach this same conclusion even if one were convinced that there were *no actual cases* of either purely competitive or purely monopolistic market conditions in economic life. Therefore, the chapters which follow will present the theory of price determination under competition and monopoly as well as under monopolistic competition, and the theory of monopolistic competition will be neither emphasized nor slighted.

QUESTIONS AND PROBLEMS

1. "Our modern economic system, with its tremendous volume of exchanges, could not have developed on the basis of barter." Do you agree? Explain.

2. "The difficulties of direct exchange are readily eliminated when money and credit are introduced into the exchange process." Explain.

3. Distinguish between value and price.

4. "Value and price may amount to approximately the same thing at a given time, but they may actually move in opposite directions over any considerable period of time." Show whether you agree.

5. "The term 'market,' as used in the theory of value, is rather indefinite." Explain.

6. Explain fully what we mean by competitive conditions in the market for an economic good.

7. Select a certain economic good and try to decide whether or not market conditions are competitive in connection with its purchase or sale.

8. "Conditions of monopoly are even more uncommon than conditions of competition in the markets of our economic system." Do you agree? Explain.

9. "A monopolist is likely to be much more powerful than a competitive producer, but his objective in carrying on his enterprise is the same as that of the competitive producer." Discuss.

10. Why do we argue that patents and copyrights are more likely to involve their owners in monopolistic competition than to give them outright monopolies?

11. Distinguish between duopoly and oligopoly.

12. "Many apparent cases of oligopoly turn out in practice to be monopoly or a form of monopolistic competition." Explain.

13. "The chief difference between conditions of competition and those of monopolistic competition is found in the number of sellers who offer a certain economic good." Do you agree? Explain.

14. What is meant by "product differentiation"?

15. "Smoking tobacco contained in a glass jar is significantly different from the same tobacco in a paper package." Show whether you agree.

16. "The economic position of the monopolistic competitor is somewhat weaker than that of the monopolist." Explain.

17. "In the markets of our economic system, conditions of monopolistic competition are more common than those of competition or of monopoly." Show whether you agree.

See the References for Further Reading at the end of Chapter XVI.

5. "Since each family requires only one house to live in whether houses are high or low in price, it is clear that the total demand for houses in a given market may not be expected to obey the Law of Demand." Discuss.

6. "The Law of Diminishing Utility states that the gratification or satisfaction derived from a unit of an economic good tends to decline as successive units of the good are consumed." Do you agree? Explain.

7. State the Law of Diminishing Utility and illustrate it from your own experience.

8. "The fact that a second unit of an economic good may actually give more satisfaction than the first unit indicates that there may be increasing as well as diminishing utility." Show whether you agree.

9. Distinguish between demand and market demand.

10. "A decrease in demand will cause the price of an economic good to fall." "A fall in price will cause the demand for an economic good to increase." Are these statements really contradictory? If not, how do you explain the apparent conflict between them?

11. The federal tax on automobiles is really no burden to the car owner. When the tax was imposed, it undoubtedly increased the prices of automobiles, but the increases in prices caused the demand to decrease, and the falling off in demand brought about lowered prices, so the prices eventually came to be as low as they were before the tax was imposed." Show whether you agree.

12. Distinguish between the two possible meanings of the term "increase in demand."

13. What is meant by elastic, inelastic, and unit elastic demand?

14. "No demand schedule is likely to be entirely elastic or entirely inelastic." Explain.

15. "The demands for some products are predominantly elastic while those for other products are predominantly inelastic." Explain.

16. What assumption concerning the nature of the demand for railroad passenger service must the Interstate Commerce Commission have made in deciding that a reduction in the basic fare from 3.6 to 2 cents per mile would help the railroads financially? Explain.

17. How does the demand for the product of one seller change as one passes from competition to monopoly and monopolistic competition?

18. Distinguish between price (average revenue) and marginal revenue.

19. "Price and marginal revenue are the same for the competitive seller but quite different for the monopolist or monopolistic competitor." Show whether you agree.

20. "Marginal revenue can be positive only when demand is elastic." Explain.

See References for Further Reading at the end of Chapter XVI.

XII

Supply and Cost of Production

The Definition of Supply. In the present chapter, we shall be concerned with the nature of supply and cost of production and with their relationship to each other and to price. Interactions of supply and demand, and the actual determination of prices, will be left for discussion in the chapters which follow. Supply, like demand, is defined in terms of schedules of quantities and prices. That is, *supply is a schedule of the quantities of an economic good which sellers would be willing to offer at a corresponding schedule of prices in a given market at a given time.* However, we must note at once that supply in this sense exists only under competitive conditions of the market. Under conditions of monopoly and monopolistic competition there is no supply schedule for an economic good in any period of time. In the intermediate and long-run periods there are cost schedules under monopoly and monopolistic competition which show the cost per unit at which various amounts of the economic good could be turned out, but the monopolist or monopolistic competitor is never willing to offer various amounts of an economic good for sale at various prices under ordinary conditions of demand. Instead the only amount of an economic good which the monopolist or monopolistic competitor is willing to offer for sale is the amount which is expected to produce the most favorable financial result in the period of time in question. Since no other amount would be offered for sale under monopoly or monopolistic competition, there is only a market supply, or a specific amount which will be offered for sale at a specified price, under these conditions of the market.

To put it in another way, the individual competitive seller is too small to influence the price of an economic good by his actions. He must therefore regard price as an objective phenomenon to which he can only react by selling or not selling his good. The total supply schedule in a competitive market is simply the result of the reactions of all the sellers to the prices which may prevail for the good. The monopolist or monopolistic competitor can change the price of his good in the market by changing the quantity of it which he offers for sale. His decision to sell a certain amount

of the good therefore cannot be interpreted as a simple reaction to an established price in the market, and it can be made only with reference to the demand which prevails for the good. On the basis of demand conditions, then, he determines to sell that market supply which he thinks will give him maximum profit or minimum loss.

SUPPLY IN THE SHORT RUN

The Fixed Stock. Whether supply must be interpreted as a series of quantities or as only a single quantity of an economic good under given conditions of the market, one of the chief considerations to occupy us at present is the fact that the amount or amounts to be supplied must be derived from different sources in different periods of time. In the short run, which is usually taken to mean a period of time too short for a changed rate of production of a commodity to be attained and brought to bear on the market, the amount or amounts to be offered for sale must obviously come from what is called a fixed stock of the good in question. *This fixed stock is a definite amount of a good already produced, on hand, and ready for sale. It is an amount which cannot be increased in the short run and can be decreased only by sale* (if we disregard fire, flood, and other calamities). For example, the retail demand for granulated cane sugar in our town of X will be satisfied by the sellers of the community on a given day out of the fixed stock of sugar which these sellers are holding. If the demand is much greater than it is expected to be on the given day, or much smaller, sellers can respond only by selling larger or smaller amounts than they expected to sell out of their fixed stock. The period of time in question is too short to permit sugar refineries to speed up their rate of operation and get deliveries to the sellers, and it is much too short to allow new refineries to be built and new machinery to be manufactured and installed. It is also too short, of course, for the sugar refining industry to reduce either its rate of operation or its productive capacity.

When World War II began in Europe in September, 1939, many consumers of sugar, remembering the high prices and limited supplies of sugar with which they had been troubled during World War I, hastened to grocery stores to lay in one- and two-hundred-pound sacks instead of their usual five- or ten-pound lots. Grocers could sell them only the amount of sugar which was on hand, and then stop selling. It was not unusual to hear a grocer say that he had sold out in the first day an entire delivery of sugar and would not receive another for a week. Somewhat similar developments occurred later in the war period in connection with sugar, coffee, butter, meats, canned goods, and other commodities, when shortages and rationing were in prospect. A large increase in the demand for a good in the short run tends to drive the price of the good up, but it can be met

only by selling large amounts out of the quantity which sellers have on hand at the time.

In the case of an agricultural commodity, such as raw cotton, new stocks become available only as additional crops come in at considerable intervals of time, and it is clear that, in the short run in a given market, sellers can respond to changes in demand and prices only by parting with larger or smaller amounts of cotton out of the fixed stock which they are holding. However, the sellers of a manufactured commodity are usually just as dependent on a fixed stock, at a given time in a given market, as the sellers of agricultural commodities. A sudden increase in the demand for shoes at retail could be met by sellers only out of the stock on hand. Some little time would have to elapse before new production at the factories could affect the retail market, and a much longer period of time would be required before new productive facilities could be brought into shoe manufacturing and set to work. The fact that the supply of an economic good is dependent upon a fixed stock in the short run is very important in connection with price determination, since this dependence limits the ability of supply to respond to changes in demand in such a period. Supply, or the amount supplied, depends upon a fixed stock in the short run, whether conditions on the supply side of the market are those of competition, monopoly, or monopolistic competition.

Supply and Cost under Short-Run Competitive Conditions. When competitive sellers are reaching decisions as to the disposal of their fixed stocks of an economic good in the short run, the cost of production of the good is not a very important consideration. Once a good has been produced and is in the hands of sellers, with costs of production already expended, the only thing that can be done under competition is to sell the good sooner or later for whatever price can be obtained, whether this price is above or below the cost of production per unit for the good. Of course, the inability of sellers to get a price for a good in the short run which will cover its cost of production per unit, or their ability to get a price in the short run which will much more than cover its cost of production per unit, will have a marked influence on their willingness to produce and sell the good over a longer period of time, but for the moment their hands are tied. The goods on hand have been produced, their costs of production are already gone, the goods cannot be held indefinitely, and their relatively early sale is indicated at the best price which can be obtained.

The Supply of the Individual Competitive Seller. While the historical costs expended on a good are of little importance to the competitive seller of an economic good in the short run, there is another interpretation of cost which is important to him. In other words, the real cost (opportunity cost) per unit in selling his stock of the good now is the price per unit which could be obtained for the good by holding it until later on, for by

selling the good now he gives up his opportunity to obtain the future price for it. Consequently, the individual competitive seller will decide to sell or hold his stock of a good on the basis of a comparison of the present price and expected future price. But the present price cannot be compared directly with the expected future price, for holding the good until later on will involve some actual or possible additional expenses. The seller must consider such factors as the cost of storing the good until some future market day, the possibility that the good may deteriorate either physically or economically before he can get to sell it in some future market, the loss of interest which will occur on funds tied up in the good if he holds it for a later market day, and the extent of his present need for cash.

When such factors have been allowed for, the future price may be spoken of as "discounted" and may be compared with the price obtainable at present. The behavior of the individual competitive seller is then easy to predict. If the present price is equal to the discounted future price, there is no reason to hold the good until later on and the individual competitive seller will sell all his stock now. If the price obtainable at present exceeds the discounted future price, he will not only sell his stock on hand but may, if the market for his good permits, sell additional quantities short at the present price for future delivery. Finally, if the price obtainable at present is lower than the discounted future price, he will hold his stock for a future market day and may even buy up additional amounts for holding, if possible.

The Supply Schedule for the Entire Competitive Market. On the basis of this supply-behavior of the individual competitive seller, it might seem difficult to develop a supply schedule for all competitive sellers of a good in the short run if the supply schedule must show that different quantities of the good would be offered for sale at different prices in a given market at a given time. That is, one might think that, at a given present price, the sellers would either sell all they had of the good or else would not sell any. This conclusion, however, overlooks important differences between the individual competitive sellers. At any given time, some sellers will expect the discounted future price to be higher than the present price of the good, while others will expect it to be the same or lower. Moreover, some sellers are conservative while others are willing to take a chance. In all probability, then, a given present price for the good will cause some sellers to hold their stocks and perhaps try to buy more while other sellers will sell all they have and perhaps even sell short.

In this situation one thing may be predicted with certainty. The higher the present price of the good, the smaller will be the number of sellers who will expect the discounted future price to be still higher, and the greater will be the number of sellers who will dispose of their stocks at present. Conversely, the lower the present price of the good the greater will be the

number of sellers who will expect the discounted future price to be higher and the smaller will be the number of sellers who will dispose of their stocks at present. These conclusions form the short-run basis for the Law of Supply, which holds that *the quantity of an economic good which sellers are willing to offer varies directly with the price in a given market at a given time.*

A possible short-run supply schedule for granulated cane sugar in our competitive market of X is presented in Table 12. In this schedule, we note

Table 12: The Supply of Granulated Cane Sugar in the Market of X on a Given Day

Price Per Pound	Quantities Which Sellers Would Offer (*in pounds*)
10¢	3,000
9¢	2,800
8¢	2,600
7¢	2,300
6¢	2,000
5¢	1,700
4¢	1,300
3¢	900
2¢	400

that, at the unusually low price of 3 cents per pound, only 900 pounds would be thrown on the market by the competitive sellers. At the usual price of something like 6 cents per pound, they would be willing to part with 2,000 pounds of sugar out of their holdings, while at the high price of 10 cents per pound they would gladly sell 3,000 pounds on this day. If the price were high enough, all the sellers would dispose of their entire fixed stocks, and in this case the market supply at that price would correspond to the total fixed stock. The data presented in this table, like those previously given for demand, are hypothetical in character and serve merely to illustrate the general nature of the short-run supply for an economic good under competitive conditions.

Increases and Decreases in Supply. Increases and decreases in supply are defined in about the same way as increases and decreases in demand. An increase in supply means that, because of some change in actual or anticipated market conditions, more sellers than formerly become willing to part with their fixed stocks at each price which may prevail in the present market. A decrease in supply, on the other hand, means that fewer sellers than formerly are willing to dispose of their fixed stocks at each price which may prevail in the present market. In Table 13, the original supply for granulated cane sugar in the market of X is reproduced from

Table 13: Original, Increased, and Decreased Supply of Granulated Cane Sugar in the Market of X

Price Per Pound	Quantities Which Sellers Would Offer Originally (*in pounds*)	Quantities Which Sellers Might Offer if Supply Increased (*in pounds*)	Quantities Which Sellers Might Offer if Supply Decreased (*in pounds*)
10¢	3,000	4,500	1,500
9¢	2,800	4,200	1,400
8¢	2,600	3,900	1,300
7¢	2,300	3,450	1,150
6¢	2,000	3,000	1,000
5¢	1,700	2,550	850
4¢	1,300	1,950	650
3¢	900	1,350	450
2¢	400	600	200

Table 12, and additional data are given which represent an increase and a decrease in this supply. Supply curves which represent the original, increased, and decreased supply of sugar in this market on this day are shown

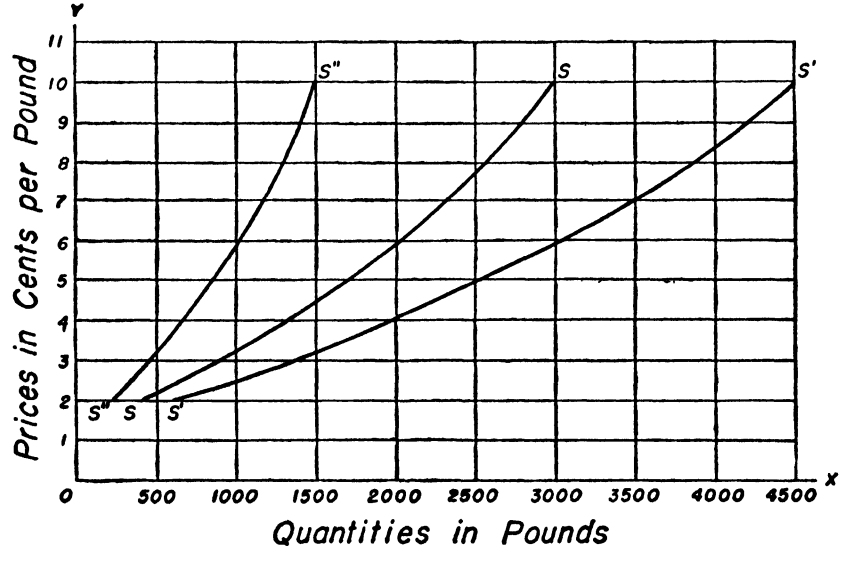

FIGURE 10.—Original, Increased, and Decreased Supply of Granulated Cane Sugar in the Market of X

in Figure 10. The curve labeled *S* represents the original supply, *S'* indicates an increase in supply, and *S"* shows a decrease in supply. It will be noted that an increase in supply is indicated by a new supply curve which, at all prices, is further to the right than the old one, while a curve representing

decreased supply is further to the left than the old curve at all prices. Since the amount of a good supplied by a monopolist or a monopolistic competitor depends upon the demand for the good and cannot be interpreted as a simple reaction to price, it will be discussed later in connection with actual price determination.

COST OF PRODUCTION AND SUPPLY IN THE INTERMEDIATE PERIOD

In the short run, as we have seen, the supply schedule for an economic good under competition and the amount supplied under monopoly or monopolistic competition are based upon a fixed stock of the good in question. In this period, moreover, the actual historical cost per unit of producing the good is of comparatively slight significance. However, in the longer periods of price determination (intermediate and long run), goods have not already been produced nor have all costs been incurred, and decisions as to the rate of production which should be maintained are conditioned to a great extent by the relationship which exists between prospective selling price and cost of production.

Opportunity or Alternative Costs. Before dealing more specifically with cost of production in the intermediate period, we must investigate briefly the general nature of cost of production. On the surface, cost of production appears to include a great variety of items, such as freight, lawyers' fees, depreciation, insurance, taxes, advertising, heat and light, stationery, postage, and a host of other things. However, upon analysis, the various things which enter into cost may all be resolved, directly or indirectly, into payments for the agents of production. Fundamentally, then, cost of production means the payments which must be made for land, capital, and labor (including managerial ability). And, since the means of production are scarce in relation to the ends to which we should like to devote them, these payments for the agents of production can be explained most satisfactorily in terms of opportunity costs.

Since most agents of production, or grades of the agents of production, are capable of being used by many enterprises in a number of different industries, the cost of using certain units of an agent to produce a given product is the value of the other products which the same units of the agent could otherwise have produced. The cost of using a certain piece of land to produce corn is the value of the soybeans or wheat which could have been produced on it, and the cost of using certain workers to produce airplane engines is the value of the automobile engines which these workers could have produced. The various enterprises in different industries bid for the quantities of the agents of production which they desire and, on the basis of these competitive bids, the supply of each agent of production (or grade of an agent) is distributed among the various enterprises and industries wishing to use it. The end result, under competition, is that an

agent of production tends to command the same price in each of its various alternative uses, though the different industries may use widely varying amounts of the agent.

If land of a certain type has a value-product of $15 per acre per year in producing corn, but one of $20 in producing soybeans, it is profitable to shift some of this land from corn production to soybean production. This process of shifting will tend to stop only when the land has the same net value-product per unit per year in both uses at some figure between $15 and $20. Hence we say that, if enterprisers in a given industry desire to use a certain quantity of a productive agent, they will have to pay as much per unit for it as anyone else is willing to pay. That is, they will have to pay the owners of these units of the agent enough to make them willing to give up their other opportunities for the remunerative employment of the units of the agent in question. This requires paying as much as the units of the agent could command in their alternative uses. Hence the name *opportunity* or *alternative costs.*

Our discussion of opportunity costs has assumed thus far that an agent of production is capable of being used in two or more alternative lines of production. If an agent of production could be used in only one industry, it would have no cost (opportunity cost) to that industry as a whole (all firms considered together) since its use in that industry would involve no foregone alternatives. However, the price per unit of the agent in question would be a cost to individual competing firms in the industry; that is, the cost of a unit of the agent to one firm would be the amount this unit would be worth to other firms.

A related case is that in which an agent of production has a high value-product in one industry but a low one in other industries. If prairie land outside a city is worth $50 per unit per year as a site for an airport but only $10 as corn land, the cost (opportunity cost) of the land to the industry which produces air transportation is only $10 per unit per year and the other $40 is noncost outlay. However, the entire $50 is cost to an individual firm producing air transportation in competition with other firms. In this situation we might grow old and gray waiting for enough of this land to be shifted from corn production to the production of air transportation so that it would command the same price in each of these alternative uses. In the more common situation in which an agent of production has two or more closely competing uses, it does tend to command the same price per unit in each of these alternative uses in the long run under competition.

The Intermediate Period. With this brief introduction, we turn to our analysis of cost of production in the intermediate period. In this period the amount or amounts of an economic good to be supplied are not based upon a fixed stock of the good in the hands of sellers. Instead, supply in this

period is based upon the rate of production which the firms producing the good can achieve with the fixed productive facilities (plant, machinery, and equipment) already at their disposal. The intermediate period is not long enough for new factories to be built or new machines and equipment to be produced and installed, but it is long enough to enable the firms in the industry to vary their rate of production considerably. They can cut their rate of production in half or even stop producing altogether, or they can expand their rate of production up to the maximum permitted by the available plant, machinery, and equipment in the industry. The situation of the individual firm in the intermediate period, with some agents of production fixed in amount and others variable, is the same situation with which we dealt in some detail in Chapter VI in connection with the Law of Diminishing Productivity.

Fixed and Variable Costs. Fixed and variable productive agents give rise to fixed and variable costs in the intermediate period. By fixed costs we mean costs of production which remain constant in total amount whether the firm operates at full capacity, near capacity, or a small fraction of capacity. While fixed costs remain constant in total amount, fixed cost per unit of product (or average fixed cost) declines steadily as output increases. Fixed costs include such items as interest payable on capital funds borrowed for investment in plant and equipment (say by means of a bond issue), rent payable on a long-term contractual basis to persons who furnish the firm with land, fire insurance on factories, salaries of certain executives or officials of the firm, retainer fees paid to lawyers, and other items. Variable costs are those which vary in total amount directly with the rate of operation or output achieved by the firm. While total variable costs vary directly with output, they do not vary proportionately with output. If they did, variable cost per unit of product (average variable cost) would be constant, and this would constitute a denial of the Law of Diminishing Productivity. Variable costs include such items as the amounts paid for raw materials and supplies, for the wages of ordinary workers, and for power used in the factory.

Average variable cost at any output is total variable cost divided by output. The typical behavior of average variable cost may be seen in Figure 11.[1] As the curve *AVC* indicates, average variable cost declines for a time as output increases. However, at still larger outputs, average variable cost increases. This behavior of average variable cost is not the result of any changes in the quality or price of the variable productive agents, for these factors are assumed to remain constant at all stages of output. The decline in average variable cost in the early stages of increasing output is quite simply the result of operating under conditions of increasing productivity,

[1] For the computation of average fixed, average variable, average, and marginal cost, see the discussion in Chapter VI.

where the quantity of the variable agents in use is relatively insufficient with respect to the fixed agents. In this stage, a given increase in output requires a less than proportionate increase in the quantity (and total cost) of the variable agents. As output is increased further, average variable

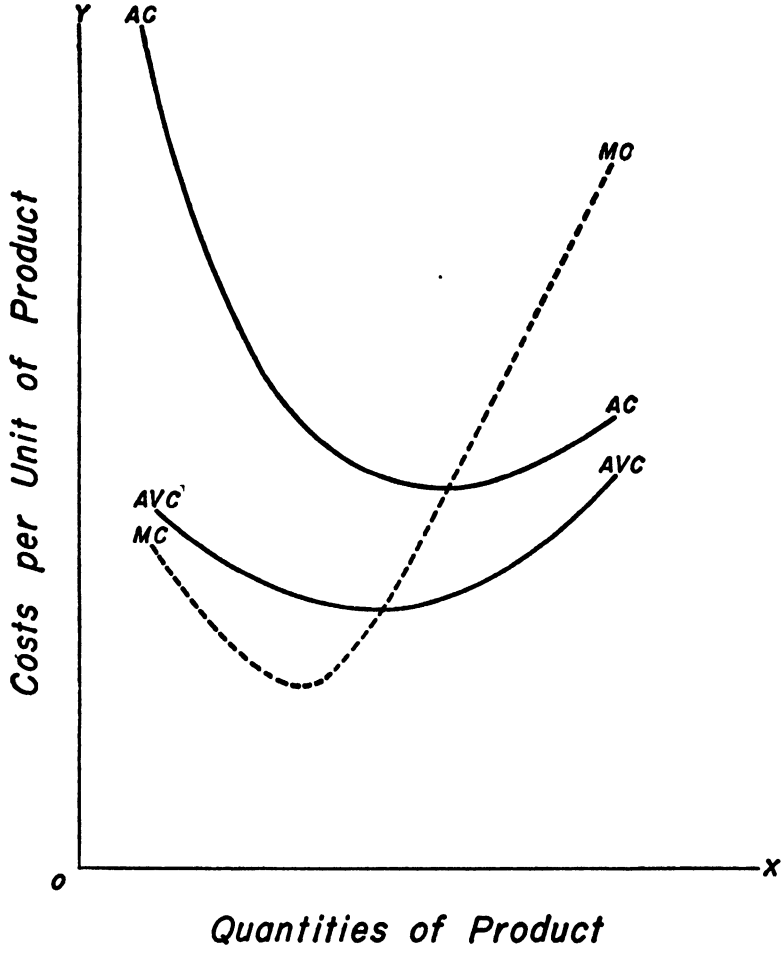

FIGURE 11.—Average Variable, Average, and Marginal Costs for the Individual Firm in the Intermediate Period

cost rises because the stage of diminishing productivity is reached and the quantity of the variable agents of production in use has become relatively oversufficient with respect to the fixed agents. In this stage, a given increase in output requires a more than proportionate increase in the quantity (and total cost) of the variable agents. Average variable cost is far below average

cost at small outputs when fixed cost per unit of output is large but approaches average cost at the larger outputs when fixed cost per unit of output is small. The low point of average variable cost occurs at the output at which most efficient use is made of the variable agents of production, and it corresponds to the point of diminishing productivity which was discussed in Chapter VI.

Average Cost. Average cost of production at any output means, of course, total costs divided by the number of units produced. As indicated by the *AC* curve in Figure 11, average cost is very high at small outputs because the burden of fixed costs is spread over only a few units of product. As output increases, average cost falls rapidly both because fixed costs are spread over more units of product and because average variable cost declines to some extent. Average cost may continue to fall for a while after average variable cost starts to increase, for the decline in fixed cost per unit of product may more than cancel the increase in average variable cost. Eventually, however, as output increases, the increase in average variable cost per unit of product more than offsets the decline in fixed cost, and average cost rises. The lowest point of average cost represents the output at which the fixed and variable agents of production are used most effectively, and this output, as we noted in Chapter VI, is often called normal or optimum output.

Marginal Cost. Marginal cost means any increase in total costs divided by the corresponding increase in output. In other words, it is the cost per unit of adding a certain number of units of product to output. As long as each unit of variable productive agents contributes more to total product than did the preceding unit (marginal product is increasing), marginal cost must fall. On the other hand, whenever each unit of the variable productive agents contributes less to total product than did the preceding unit (marginal product is declining), marginal cost must rise. Marginal cost is equal to average variable cost at the lowest point of average variable cost, and to average cost at the lowest point of average cost.[2] The lowest point of

[2] The relations of marginal cost to other types of cost may be explained readily. Marginal cost is the average variable cost of adding a certain number of units to output, while average variable cost is the total variable cost of a certain output divided by that output. As long as the cost per unit of adding units to output (marginal cost) is below the average variable cost of all previous units of output, average variable cost itself must fall. This is true whether marginal cost is falling or rising, just so long as it is below the level of average variable cost. On the other hand, when the cost per unit of adding units to output (marginal cost) is above the average variable cost of all previous units of output, average variable cost itself must rise. Since average variable cost must fall as long as it is above marginal cost and must rise whenever it is below marginal cost, average variable cost will be neither rising nor falling (that is, it will be at its lowest point) when it is equal to marginal cost.

The relation of marginal cost to average cost is similar. As long as the cost per unit of adding units to output (marginal cost) is below the average cost of all previous units, average cost itself must fall (just as the average weight of a group of men would fall if one already had three men who averaged 150 pounds and added another who weighed only 100 pounds). Conversely, when the cost per unit of adding units to output (marginal cost) is above the average cost of all previous units of output, average cost itself must rise (just as the average

marginal cost itself has no significance for our analysis. It would correspond to the point of diminishing productivity if this latter point were defined in terms of marginal product rather than in terms of average product per unit of the variable agents.

Supply of the Individual Competitive Seller. The preceding cost analysis applies to any individual firm, whether it operates under conditions of competition, monopoly, or monopolistic competition, but the analysis of supply can be developed at this point only for competitive conditions of the market. As we have said, the monopolist or monopolistic competitor is able to influence the price of a good by varying the amount supplied, so that the amount supplied by such a seller cannot be explained as a simple reaction to price in the market. However, the individual competitive seller is willing to supply various amounts of the good at the different prices which might prevail in the market, and we must now analyze his supply situation.

For this purpose, marginal cost is more important than the other types of cost which we have studied. The general principle of supply for the individual competitive seller is that he will add to his output and amount supplied as long as each unit of product adds more to his income than to his costs. Price and marginal revenue, it will be recalled, are the same thing for the individual competitive seller, and so, when a unit of product is sold at a certain price, the amount of the price is added to the seller's income. On the other hand, the amount a unit of product adds to the seller's costs is the marginal cost of the unit. It is tempting, therefore, to say that the entire marginal cost curve of the individual competitive firm is its supply curve in the intermediate period, and that the firm will supply any amount of the good at a price which covers marginal cost for that output.

However, this conclusion is subject to one important qualification. It would be possible, as a glance at Figure 11 will show, for price to cover marginal cost without covering average variable cost. Now variable costs are those which are contingent upon output and could be eliminated entirely if the enterprise temporarily ceased to operate. Hence there is no point in producing any output, even though the price obtainable will cover marginal cost, if the price does not also cover average variable cost. Our conclusion is that the marginal cost curve of the individual competitive firm is its supply curve but only in the area from the intersection of marginal cost with average variable cost to the upper end of the marginal cost curve. Within this area, the individual competitive firm will expand output and amount supplied as much as possible without having marginal cost exceed

weight of a group of men would rise if one already had six men who averaged 135 pounds and added another who weighed 200 pounds). Since average cost must fall as long as it is higher than marginal cost and must rise as long as it is lower than marginal cost, average cost will be neither rising nor falling (that is, it will be at its lowest point) when it is equal to marginal cost.

price. In other words, within this area the firm will supply any output of the good at a price which covers marginal cost. Since the marginal cost curve, within the area in which it is a supply curve for the individual competitive firm, moves steadily upward with increased output, it is clear that the amount supplied by the firm in the intermediate period will vary directly with the price obtainable for the good in the market.

If the output at which price is equal to marginal cost is greater than normal output, a price equal to marginal cost will more than cover average cost, and the individual competitive firm will be operating at a profit. This, of course, is quite possible in the intermediate period. On the other hand, if the output at which price is equal to marginal cost is smaller than normal output, a price equal to marginal cost will fail to cover average cost, and the firm will be operating at a loss. However, since such a price will more than cover average variable cost, the receipts of the firm will cover its variable costs and part of its fixed costs, whereas, if it did not operate at all, its fixed costs would be a total loss. Thus the firm would be better off (in the intermediate period only) to produce and sell at a loss than to close down, if the price is equal to marginal cost and greater than average variable cost.

Supply of a Whole Competitive Industry. The next question is how we can pass from the supply curve of a single competitive firm to the supply curve of a whole competitive industry in the intermediate period. It is obvious that in a competitive industry in the intermediate period, with its large number of small producers, it would be a miracle if all producers had the same fixed, variable, average, and marginal costs of producing a given economic good. All firms would have all of these types of costs, of course, but they would be at different levels for different firms and would have different rates of increase and decrease. Actually the solution of this problem is comparatively simple. Since the supply curve of each competitive firm is its marginal cost above the level of average variable cost, the supply curve for a whole competitive industry can be found by adding together horizontally the marginal cost curves (above the level of average variable costs) of the individual firms.

Let us see what this means. Let us say that, at a price of 80 cents, the firm whose costs are illustrated in Figure 11 could produce 8,000 units without having marginal cost exceed price, but other firms could produce only 6,000 or 7,000 units at this price while still others could produce 9,000 or 10,000. By adding together the outputs which all firms could afford to produce at a price of 80 cents per unit, we could obtain the amount which the whole industry could supply at this price. The same procedure would have to be followed for other prices which might prevail for this good in the market. When the task of adding the outputs of individual firms at all prices had been completed, we would have a supply schedule for the whole

competitive industry in the intermediate period, and this schedule would show the amounts of the good which the whole industry would produce and offer for sale at all probable prices.

Since the marginal cost curve of each individual firm, above the level of average variable cost, rises steadily with increases in output, it is clear that the same thing would be true of the curve obtained by adding together the marginal cost curves of the individual firms. In other words, the supply schedule for a whole competitive industry in the intermediate period

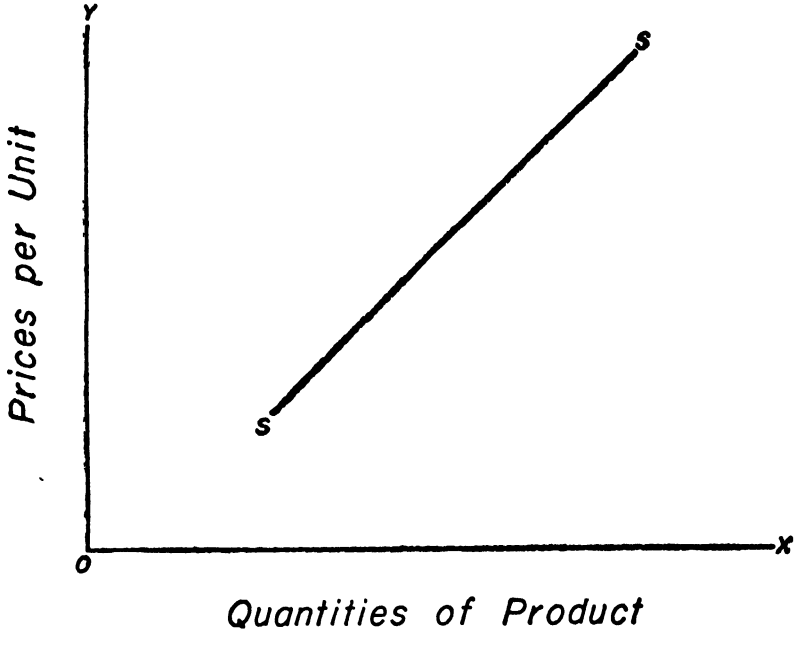

FIGURE 12.—The Supply Curve for a Whole Competitive Industry in the Intermediate Period

would show that the industry would produce and offer for sale larger amounts of the good at high prices than at low prices, in conformity with the Law of Supply, and the supply curve representing this supply schedule would take the form of the curve *S* in Figure 12 (although the curve would not neeessarily be a straight line). While all competitive firms in the industry would receive the same price for the good and, in the case of each firm, this price would be equal to marginal cost, it is quite possible that some firms would be operating at a rate in excess of normal output while others would fall short of normal output at the same time. In other words, the same price would be likely to bring profits to some firms and losses to others.

COST OF PRODUCTION AND SUPPLY IN THE LONG RUN

The Long Run. Finally, we come to the subject of supply and cost of production in the long run. In this period the amount or amounts of an economic good to be supplied are still based upon the rate of production achieved by the industry in question, but the rate of production can now be varied indefinitely. This means that the long run is long enough to permit additional quantities of those agents of production, which were fixed in amount in the intermediate period, to be brought into the industry and placed in operation, or to permit some of these agents of production to pass out of the industry, if a contraction in productive capacity, instead of an expansion, is indicated by demand conditions. In the long run, changes in the rate of production are the result of changes in the size and productive capacity of an industry and are not accomplished by running faster or slower than before on the basis of existing productive facilities. The basis for supply or amount supplied in the long run is the same under competition, monopoly, or monopolistic competition, for under each of these conditions the long run is a period of time in which the productive capacity of an industry can be increased or decreased.[3]

Changes in Industrial Capacity. In studying the effects of changes in the productive capacity of an industry upon cost of production, it is important to know how these changes in productive capacity may be expected to take place. Will the expansion or contraction of the industry also involve increases or decreases in the size of the individual firms in the industry, or will it be accomplished by the entry of new firms into the industry or the exit of some old firms from the industry? If changes in the size and productive capacity of an industry involve changes in the size of the individual firms in the industry, we already know something of what to expect, in

[3] Now that we have examined all three of the periods of time used in the study of price determination, we should emphasize that it is not possible to define these periods exactly in terms of clock or calendar time for all industries. Instead, they are defined in terms of "operational time"; that is, each period is long enough to permit the accomplishment of certain results. The long run, we have said, is a period sufficiently long so that the productive capacity of an industry can be increased or decreased by bringing new productive facilities (plant and equipment) into the industry or by allowing such productive facilities to drop out of use. In some branches of agriculture such a change in productive capacity might be accomplished from one year to the next. In some manufacturing industries a period of two or three years might be required. In the production of raw rubber or bananas, where several years must elapse before new plants can be made to yield the desired product, the long run would be a very extended period. The intermediate period is also of varying length in terms of clock time. It is long enough to change the rate of production in an industry but not long enough to change its productive capacity. In some branches of agriculture it may be several months or even a year before a change can be made in the rate at which existing productive facilities are utilized. In some manufacturing industries only a few days may be needed to speed up or slow down production. The short run is a period of time in which changes in the rate of production cannot be brought to bear upon the stock of a good on the market, but this may mean anything from a few days up to several weeks or months. Each of these three periods, therefore, should be thought of purely in terms of the basis upon which the supply of an economic good rests in that period, without regard to the actual time content.

regard to the effect on cost of production, from our discussion in Chapter VII on the size of the firm in relation to cost.[4]

As the individual firm expands from small size to what has been called optimum size, its average costs descend to lower levels. This results from the fact that, at optimum size, the firm is able to make most effective use of a number of indivisible productive agents and services which could not be used effectively when the firm was at smaller size. The result, in terms of average cost curves, is seen in the transition from AC^1 (in Figure 13),

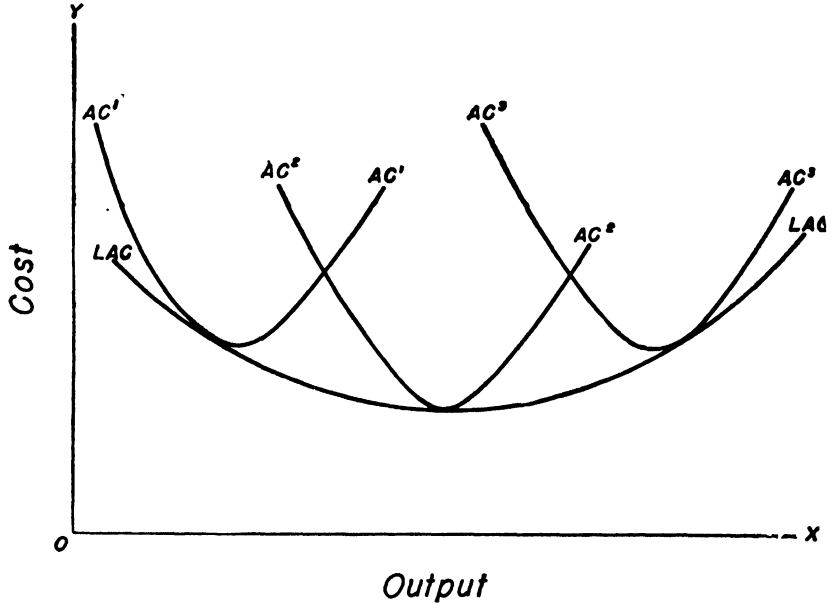

FIGURE 13.—Average Cost Curves in Relation to Size of Firm, and Long-Run Average Cost Curve for the Individual Firm

which represents a small size of the firm, to AC^2, which represents the optimum size of the firm. However, as the firm expands beyond optimum size, any further economies of growth are more than completely offset by diseconomies which result from the grave difficulties of large-scale management, and its average costs ascend to higher levels. The result, in terms of average cost curves, is seen in Figure 13 in the transition from AC^2, which represents the optimum size of the firm, to AC^3, which represents a larger than optimum size of the firm.[5]

On the basis of this succession of intermediate-period curves of average

[4] The student should review the discussion in the early part of Chapter VII at this point.

[5] Figure 13 is not intended to imply that there are only three possible sizes for the individual firm. There may be many other possible sizes, but it would complicate the diagram unduly to introduce them.

cost representing different sizes of the firm, it is possible to draw a long-run curve of average cost for the individual firm, as shown by *LAC* in Figure 13. Long-run average cost means the lowest possible cost at which the firm can turn out each output when it has time enough to make any desired change in its size and productive capacity. On this basis, a long-run average cost curve for the individual firm must be drawn in such a way that no part of any intermediate-period average cost curve of the firm ever lies below it. If this requirement is satisfied, the long-run average cost curve will be tangent to each of the succession of intermediate-period average cost

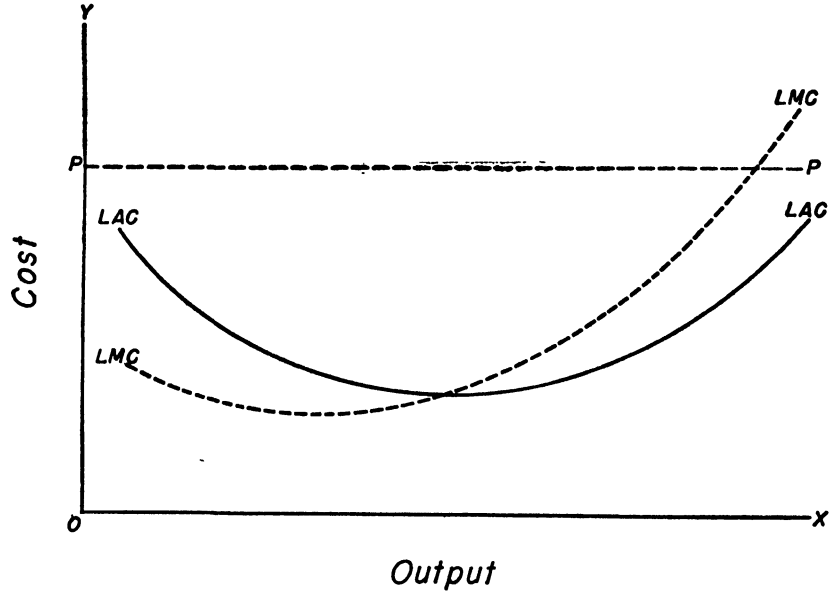

FIGURE 14.—Long-Run Average and Marginal Cost Curves for the Individual Firm

curves for the firm. It is also possible to draw a long-run marginal cost curve for the individual firm, which is derived directly from the long-run average cost curve. Such a long-run marginal cost curve (*LMC*) is shown in Figure 14 along with the corresponding long-run average cost curve. This analysis of the changing size of the individual firm in relation to cost is apparently valid whether the firm is operating under conditions of competition, monopoly, or monopolistic competition. However, for the usual reasons, we shall consider its significance only in relation to the supply reactions of the individual firm and industry under competitive conditions in the present chapter.

In a competitive industry the individual firm would show a tendency

to grow to optimum size, under any given size of the industry as a whole, on the basis of the particular productive agents which it was able to command. If the industry as a whole, starting from this situation, were to increase in size and productive capacity by means of the expansion of the individual firms, it is clear that the individual firms would have to grow to some size beyond the optimum. Is this likely to happen? If, for example, there is a great increase in the demand for the product of a competitive industry and the price rises to the level indicated by the line *P* in Figure 14, will the individual firm expand beyond its optimum size so that it can furnish the output at which its long-run marginal cost would be equal to the price? The answer is that the firm would be willing to do so, and in fact would make maximum profits by doing so, if it had any assurance that the price would remain at this level.

However, in a competitive industry, just the opposite result is assured. The price in question is well above the average cost of the firms already in the industry even though these firms expand their (intermediate-period) rate of operation to something beyond normal output. Large profits are made by the firms, and the agents of production in use in this industry earn more than those which are employed in other fields. As a result, new firms and agents of production will be attracted to the industry, its size, productive capacity, and output will be increased, and the price will fall back to its normal level equal to the average and marginal cost of the individual firm at optimum size and output.[6] In this situation, any of the original firms which had expanded to a size beyond the optimum would be at a disadvantage, with respect to cost, by comparison with new firms of optimum size. The general conclusion is, then, that the expansion of a competitive industry in the long run is more likely to occur by means of the entry of new firms than by the expansion of firms already in the industry. In similar fashion, the contraction of the size and productive capacity of such an industry is likely to involve the exit of some firms from the industry rather than a contraction of the size and productive capacity of all firms in the industry. That is, the continuation of an unfavorable price situation, resulting in losses for the firms and lower earnings for the productive agents than those received in other industries, would cause some firms and their productive agents to depart for fairer fields.

Size of the Industry in Relation to Cost. If a competitive industry in the long run expands its size and productive capacity by means of the entrance of firms, what effect, if any, is the expansion of the industry likely to have on the costs of the individual firms? In the overwhelming majority of cases, the result would be an increase in the costs of the individual firms.

[6] Whether the price will fall back to exactly its former level, or to a level slightly higher or lower, will depend upon the effect, if any, of the entry of new firms on the costs of all the firms. This matter will be investigated immediately.

If, for example, the wheat-raising industry were to increase in size and productive capacity, how would the new firms get their necessary land? They might, of course, bring into production some land which was not being used for anything at the time, but such land would not be nearly as well suited for wheat raising as that which was already in use, and land-cost per bushel of wheat would be high because of the low productivity of the land even though the land did not cost very much per unit.

The alternative to using poorer land would be to bid some high-grade land away from producers in other branches of agriculture who were using it to produce other crops. However, except to the extent that wheat was a substitute for these other agricultural products, this would amount to an increase in the total demand for the high-grade land and would result in an increase in its price or cost in all of its alternative uses, including wheat raising. The result would be that the land-cost per bushel of wheat would increase for all the firms in the wheat-raising industry. The same two possibilities of increasing cost per bushel of wheat could be developed in connection with the specific grades of labor or almost any productive agent required in wheat raising.

On the other hand, the contraction of the industry and exit of firms would be expected to lower the costs of all the firms, either because poorer grades of certain productive agents would be dropped from use or because the contraction of the industry, other things equal, would cause some decrease in the total demand for high-grade productive agents. An industry of this kind, in which an expansion of the industry tends to increase the costs of all the firms while a contraction tends to decrease the costs of all the firms, is said to operate under conditions of increasing cost in the long run.

Two other possibilities exist with respect to the effect of the expansion of a competitive industry on the costs of the individual firms. An industry would be said to operate under conditions of constant cost in the long run if neither an expansion of the industry by means of the entrance of firms nor a contraction of the industry by means of the exit of firms would have any effect on the prices of the productive agents used in the industry. Clearly this would be possible only if the industry as a whole were so small that all the firms taken together used only a minute fraction of the total available quantities of the productive agents necessary to its operation. For example, if the rubber thumb-ends used by postmen, bank clerks, and others were produced by a separate and distinct industry, the expansion of the industry and entrance of firms would not be expected to bid up the prices of any necessary productive agents nor would the contraction of the industry and exit of firms be expected to lower the prices of any of these agents. There are not likely to be many industries of constant cost in actual practice.

Finally, there is a remote possibility that the expansion of a competitive industry and the entrance of firms into it might result in the lowering of

the costs of all the firms. This might happen if the firms in the given industry obtained productive agents or services from other industries which also operated under conditions of decreasing cost. In this case the expansion of the given industry would permit the other industries to expand and operate at lower cost, thus lowering the cost of the productive agents or services to the given industry. There are obvious limits on this possibility. Again, the given industry might be purchasing from an overexpanded monopoly which was operating at less than full capacity. In this case the expansion of the given industry might allow the monopoly to increase its rate of output and secure lower cost per unit of product. The result might be a lowering of the cost of the monopoly product to the given industry. However, the monopoly, unless compelled to do so by public authority, might not pass on any such saving in cost to the given industry. In converse fashion, a decrease in the size of the given industry and the exit of firms might result in an increase in the costs of all the firms. If any industry of the kind just described actually existed, it would be said to operate under conditions of decreasing cost in the long run.[7]

The firms in an industry operating under monopolistic competition may also experience changes in their costs as the industry as a whole expands or contracts by means of the entry or exit of firms in the long run, and even a monopoly may find that its costs will change in the long run as it constructs and operates additional plants or allows old ones to pass out of existence. However, these matters can be analyzed best in connection with price determination under these conditions of the market. For the present we shall be content with a few additional observations in connection with cost and supply in the long run under competition.

Cost of Production and Supply in the Long Run under Competition. Supply and cost of production are inseparable in the long run under competition. Under any given condition of demand, the size and productive capacity of a competitive industry will be adjusted so that all of the firms in the industry, being at optimum size and producing at normal output, will turn out an amount of the product which will sell at a price equal to the average and marginal costs of the firms. And under any change in demand conditions, the size and productive capacity of the industry will also be changed until the same results are achieved. Thus it may be said that, in the long run under competition, the cost at which the firms can produce a given output is also the price at which this output will be supplied and offered for sale in the market.

The Elimination of Differential Costs between Firms. If changes in the size and productive capacity of a competitive industry in the long run have the same effect on the costs of production of *all the individual firms,*

[7] The most effective illustration of these possible long-run cost tendencies requires the use of diagrams for price determination; hence this matter will be deferred to Chapter XIV.

and if such an industry, when completely adjusted in the long run, turns out an amount of product which will sell for a price equal to the average and marginal costs of *all the firms,* the clear implication is that all the firms in a competitive industry tend to have the same costs of production in the long run. And this conclusion is justified, strange as it may seem in view of admitted differences between firms with regard to cost in shorter periods of time. The advantages which some firms have over others with respect to cost in the shorter periods of time are attributable for the most part to superior agents of production. These superior agents are either reproducible or not. If they are reproducible, all firms in a competitive industry will tend to acquire the superior agents of production in the long run. If they are nonreproducible, the value of these superior agents will tend under competition to be bid up to such a level that no net advantage in using them remains.

If some firms have superior factories, machines, or other capital goods, they will have an advantage over other firms in the short run or in the intermediate period when such productive facilities cannot be changed. In the long run, since capital goods are reproducible, all firms in a competitive industry will come to adopt these superior devices. If some firms have superior workers, they will have an advantage over other firms in the shorter periods of time, but the other firms will also acquire superior workers in the long run if more of these workers are available or can be trained. If the number of superior workers is naturally limited, the few superior workers available will find their wages bid up under competition to a point where the workers confer no net advantage on the firms which hire them. If some firms have better combinations of the agents than other firms, or better methods and processes, these combinations, methods, and processes will tend to be adopted by all firms in a competitive industry in so far as they do not depend upon limited, nonreproducible agents of production.

If the advantage which some firms have over others with respect to cost at any given time is due to land which is more fertile for agricultural purposes or better located for other productive purposes than that of other firms, the superior agent is nonreproducible. However, if the producers with the temporary cost advantage merely rent this land from its owners, they will find, since all firms will desire to use this land, that its rent in the long run under competition will be bid up to such a level that its use confers no net advantage on them. Finally, even if the firms with low costs in the shorter periods of time actually own the superior land, they will find it necessary in the long run to raise the implicit rent which they charge themselves to such a level that their cost of production will be the same as that of firms which use inferior land. This conclusion follows from the principle that, when a firm owns an agent of production, the cost of the

agent to the firm is equal to what the agent would be worth to other firms. Thus the net conclusion is that all of the many firms in a competitive industry tend to have the same minimum long-run average costs, and that this cost of production for any firm may be taken as the cost of production for the industry as a whole.

The Adequacy of a Cost-of-Production Price. We have seen that a competitive industry in the long run tends to have such a size and productive capacity and tends to produce such an output that the price of the product will be equal to the average (and marginal) costs of all firms as these firms operate at normal output. This conclusion that the firms in a competitive industry tend to receive a price equal to cost of production in the long run has had to meet a great deal of criticism. It is sometimes assumed that this statement means that each firm in the industry must sell its product for a price only high enough to cover its money expenses of production and that it will be left with no income at all for the productive agents owned and furnished by the firm itself. Thus if cost of production is taken to refer only to the out-of-pocket expenses of a firm, a price equal to cost of production would be most unsatisfactory to the owners of the firm, and a price which would exceed this cost so as to furnish the firm with a considerable margin of "profit" would be necessary if the enterprise were to continue functioning for any great length of time.

However, the economist's interpretation of cost of production is much broader than that suggested in the preceding paragraph. Cost of production includes the out-of-pocket expenses of an enterprise for wages, interest, materials, power, and so on, to be sure, but it includes much more. To these money expenses must be added the value of the enterpriser's own ability and services (in the case of the individual proprietorship or partnership), as measured by the salary which he could command if he hired himself out as a manager instead of working for himself. Again, cost of production includes interest at the competitive rate for any capital funds which the owners of the firm have furnished to the business, and rent at the competitive rate for any land which they have supplied. Since all of these agents of production furnished by the owner or owners of a firm would be able to command a price if hired out to some other business, it is ridiculous to assume that their use involves no cost to the firm whose owners have furnished them.

Average cost of production in the long run, in other words, includes payments at competitive rates for all agents of production used by a firm, whether these agents of production are owned outside the firm so that the payments for them constitute an out-of-pocket expense, or are owned by the firm (or its owners) so that the payments for them are received by the owners of the firm. Clearly, then, a price equal to average cost of production in the sense in which the economist uses that term, would not be

discouraging to the owners of a competitive firm. It would not make an enterpriser rue the day he was born or yearn to be a hired employee rather than an enterpriser. A price equal to average cost of production in the long run would be eminently satisfactory to the enterpriser in a competitive industry, since it would pay him all he is worth, and all his agents of production are worth, in connection with the business. If he could be sure of never getting less than this cost of production, he would never need to ask for more, and a profit above this cost of production, while very acceptable, would never be necessary.

QUESTIONS AND PROBLEMS

1. How do we differentiate between the short-run, long-run, and intermediate periods in connection with price determination?

2. "The definition of supply is always the same regardless of the period of time which is under discussion, but the basis upon which the supply of an economic good rests varies greatly from one period of time to another." Explain.

3. "There are supply schedules for economic goods under competitive conditions but only market supplies under conditions of monopoly or monopolistic competition." Show whether you agree.

4. What is meant by saying that the supply or market supply of a good depends upon a "fixed stock" in the short run?

5. "Opportunity costs are more important than actual historical costs when an enterpriser is trying to decide whether to sell or hold his stock of a good in the short run." Explain.

6. "A firm decides to sell or hold its stock of a good in the short run on the basis of a comparison of the present price and the discounted future price of the good." Explain.

7. "The quantity of an economic good which competitive sellers are willing to offer for sale varies directly with price in the short run." Why?

8. Summarize the theory of opportunity or alternative costs.

9. "The cost of using certain units of a productive agent to produce a given product is the value of the other products which the same units of the agent could otherwise have produced." Explain and illustrate.

10. Explain the behavior of average cost and average variable cost in the intermediate period.

11. How is the behavior of marginal cost related to that of average cost in the intermediate period? Explain.

12. "The entire marginal cost curve of the individual competitive firm is its supply curve in the intermediate period." Show whether you agree.

13. "The marginal cost curve of the individual competitive firm is a supply curve in the area above its intersection with average variable cost." Do you agree? Explain.

14. How can the supply curve for a competitive industry be derived from the supply curves of individual firms in the intermediate period?

15. "If the individual firm sells its product at a price equal to marginal cost in the intermediate period, it will sometimes make profits and sometimes take losses." Show whether you agree.

16. "If a whole competitive industry had to expand by means of increases in the size of the individual firms in the long run, we could be rather certain of what would happen to each firm's cost of production." Explain.

17. "The expansion of a competitive industry in the long run is more likely to occur by means of the entry of new firms than by the expansion of firms already in the industry." Do you agree? Explain.

18. "When a whole competitive industry expands in the long run, the costs of the individual firms are more likely to increase than to decrease or remain constant." Show whether you agree.

19. Explain fully what is meant by an industry which operates under conditions of increasing cost in the long run.

20. "All enterprisers in a competitive industry tend to have the same average cost per unit of product in the long run." Explain.

21. "Business enterprisers will not remain in business in the long run unless they can sell their goods for prices which yield a margin of profit over and above cost of production." Do you agree? Explain.

See References for Further Reading at the end of Chapter XVI.

XIII

The Determination of Prices under Competition

Having learned something of the general nature of demand, supply, and cost of production, we may now proceed to study price determination, and we begin with the determination of prices under competitive conditions of the market. In the present chapter, and the one which follows, we shall consider competitive price determination in the short run, the intermediate period, and the long run, but price determination under monopoly, monopolistic competition, or other conditions of the market will be left to later chapters.

PRICE DETERMINATION IN THE SHORT RUN

Demand and Supply Schedules. In Table 14, we bring together the demand and supply schedules for granulated cane sugar in the market of X on a given day, as presented in earlier discussions. The demand schedule, in conformity with the Law of Demand, shows that the quantity of this good which buyers would purchase varies inversely with the price. This behavior of the quantity demanded depends as usual upon differences in the money incomes and desires of individuals and the operation of the Law of Diminishing Utility. The supply schedule, as predicted by the Law of Supply, shows that the quantity of this good which sellers would offer varies directly with the price. In other words, the higher the present price, the smaller will be the number of individual sellers who will expect the discounted future price (obtainable by holding the good for a time) to be still higher and the greater will be the number of sellers who will be willing to dispose of their stocks of the good at present.

Price Determination. Since the quantity demanded varies inversely with the price while the quantity supplied varies directly with the price, it is clear that quantity demanded and quantity supplied must come together at some price or other. In the present case, as indicated by the demand and supply schedules themselves and by Figure 15, in which demand

and supply curves representing these schedules are shown, the quantity demanded and quantity supplied are equal at a price of 5 cents per pound

Table 14: Demand for and Supply of Granulated Cane Sugar in the Market of X on a Given Day

Price Per Pound	Quantities Which Buyers Would Purchase (*in pounds*)	Quantities Which Sellers Would Offer (*in pounds*)
10¢	1,000	3,000
9¢	1,100	2,800
8¢	1,200	2,600
7¢	1,350	2,300
6¢	1,500	2,000
5¢	1,700	1,700
4¢	2,000	1,300
3¢	2,500	900
2¢	3,100	400

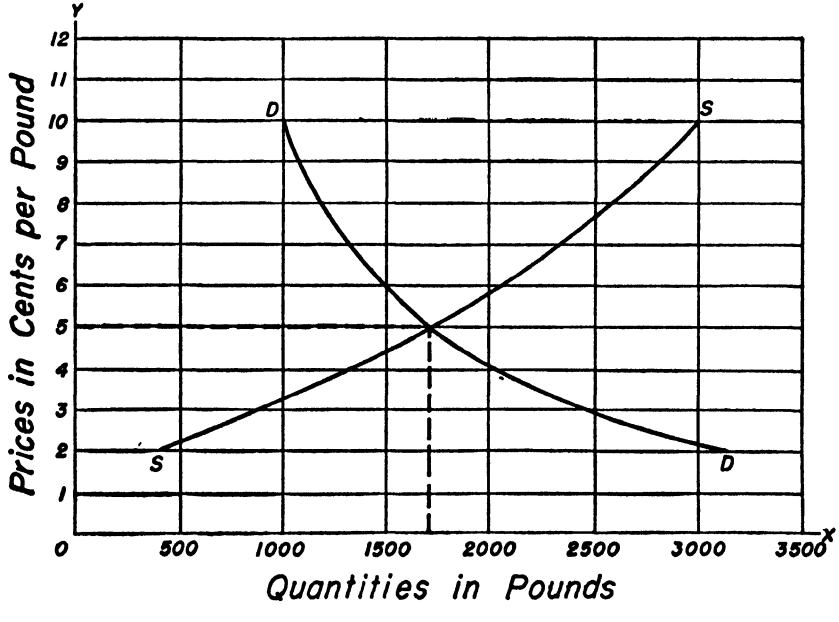

FIGURE 15.—The Determination of the Price of Granulated Cane Sugar under Competitive Conditions in the Market of X on a Given Day

for the sugar. This means that, under the assumed conditions, the price of sugar will be 5 cents per pound in the market of X on the day in question and that the quantity bought and sold will be 1,700 pounds on this

occasion. The point where the demand and supply curves intersect is known as the point of equilibrium, and 5 cents per pound is the equilibrium price, representing, as it does, a sort of balance between the opposing forces of demand and supply. Because these terms refer to the entire demand and supply schedules, it is not correct to say, as students sometimes carelessly do, that demand and supply are equal at the point of equilibrium. It is the amount demanded and the amount supplied, or market demand and supply, which are equal at the equilibrium price of 5 cents per pound. This price is a price which clears the market. No seller will have in his possession at the end of the day any stock of sugar for which he was willing to accept 5 cents per pound or less on that day and no buyer willing to pay 5 cents per pound or more for sugar on that day will go home without the sugar which he desired to purchase.

Under the assumed conditions of competition, a price higher or lower than 5 cents per pound for sugar could not prevail on this day in this market. At a price of 6 cents per pound, for example, buyers would be willing to purchase only 1,500 pounds, while sellers would like to dispose of 2,000 pounds. Some sellers who were willing to dispose of their stocks at this price would find themselves unable to do so if the price continued. However, the efforts of the sellers to do business with the more limited number of buyers would tend to bring the price down to the equilibrium level of 5 cents per pound. On the other hand, at a price of 4 cents per pound, buyers would be willing to purchase 2,000 pounds on this day while sellers would care to part with only 1,300 pounds. If such a price continued to prevail, many buyers would find themselves unable to acquire all the sugar which they desired, but, of course, their efforts to buy a large amount of sugar from the sellers at this price would tend to bring the price up to the equilibrium level of 5 cents per pound. As we noted in the preceding chapter, this equilibrium price in the short run may or may not be high enough to cover the cost per pound of producing the sugar and bringing it to the market.

The Law of One Price. Figure 15 shows all of the sugar being bought and sold in the market of X on a given day at the price of 5 cents per pound. No purchases and sales are indicated at any price higher or lower than 5 cents per pound. This situation is described by the Law of One Price, which says that, *under competitive conditions, only one price can prevail for an economic good in a given market at a given time.* Students often have difficulty in appreciating this law, for it seems to run counter to many of their experiences with actual prices. That is, in many cases different prices are apparently charged for exactly the same good in a given market at a given time. For example, the customer may purchase a large can of a given tooth powder for 35 cents at the drugstore on the corner, while by going to a chain or a cut-rate drugstore across the street he can

get the same-sized can of the same powder for 27 or 29 cents. Instances of this kind are very numerous in our actual markets.

The difficulty in such cases is, of course, found in the fact that, for some reason or other, conditions of competition are less than perfect in the market in question. In some cases there may be product differentiation even though the different sellers deal in exactly the same article. One store may be air conditioned, while the other is not, or may please the customer's eye with more elaborate appointments and furnishings. One store may give credit and furnish delivery service if desired, while the other operates on a cash-and-carry basis. The buyers may not know of the differences in prices which prevail or, knowing of them, they may not think it worth while to visit a dozen stores in order to save a couple of cents on each article purchased. The sellers in the given market may be so few that they really operate under conditions of oligopoly, and so on. If it were really true in the market that buyers and sellers of exactly the same good were very numerous, that there were no reasons for buyers to prefer one seller to another, that each buyer and seller acted independently of all others, and that both buyers and sellers were well informed concerning market conditions, only one price could prevail for the economic good in the given market at the given time. This is what the Law of One Price holds.

Cases such as those noted above, however, do not mean that the prices charged by different sellers are unrelated except where competition is perfect. It may be possible, where competition is imperfect, for one seller to get 35 cents for an economic good even though another seller in the same market charges only 29 cents for what is ostensibly the same product at the same time, but if one seller charged 1 dollar per unit and the other only 29 cents we should expect all the business to go to the latter seller even if competition were less than perfect in the market. The Law of One Price, then, might be interpreted as saying that, to the extent that conditions of competition prevail in the market for a given economic good, to that extent only one price tends to be charged for the economic good at a given time. Under perfect competition there will be only one price. Under conditions which are largely but not perfectly competitive, small differences in prices may exist but large differences cannot. If almost none of the conditions of competition are present, relatively large differences in prices may exist for the same good at the same time and in the same market.

The Law of Short-Run Competitive Price. While the short-run period is too short to permit the production of an economic good to become adjusted to the demand for it, it is possible for the demand for or supply of the good to increase or decrease in this period. That is, buyers may become willing to purchase more or less of an economic good at all prices in the demand schedule, or, because of some change in market conditions, more

or fewer sellers may become willing to dispose of their stocks at each price which might prevail for the good. Since such changes in demand or supply will tend to affect the price of an economic good, the Law of Short-Run Competitive Price is devised to cover these situations. This law states that *in a given market under short-run competitive conditions, the price of an economic good tends to vary directly with demand and inversely with supply.* That is, under the assumed conditions, if supply remains unchanged, the price of an economic good will increase when the demand for it increases and decrease when the demand decreases. On the other hand, if

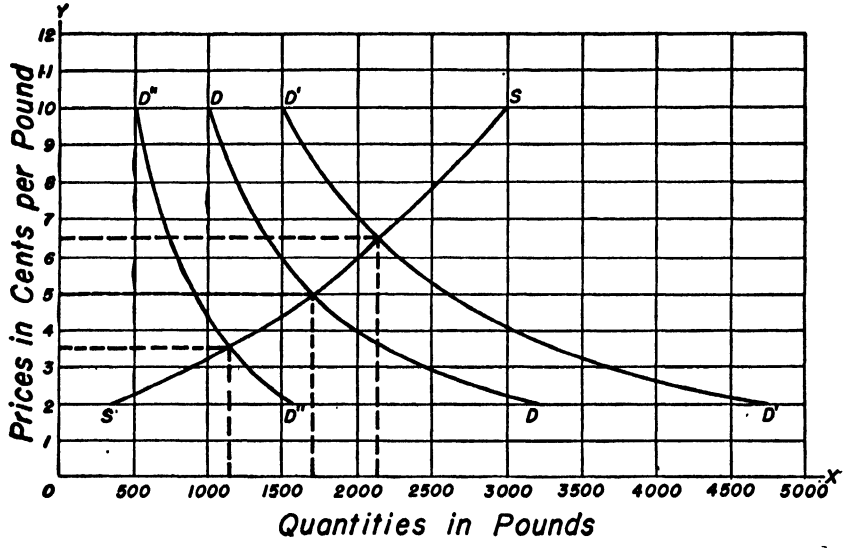

FIGURE 16.—Effect of Increased and Decreased Demand on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions

demand remains unchanged, the price of an economic good will increase when supply decreases and decrease when supply increases.

These situations are illustrated in Figures 16 and 17. The supply and demand curves in these diagrams are drawn from the schedules representing original, increased, and decreased demand for and supply of sugar in the market of X in the short run, as presented in Chapters XI and XII. In Figure 16, supply remains unchanged, and an increase in the demand for sugar from D to D' causes the price to increase from 5 cents to 6½ cents per pound, while a decrease in demand from D to D'' causes the price to decrease from 5 cents to 3½ cents per pound. The quantity bought and sold increases with the increase in demand and decreases with the decrease in demand. In Figure 17, demand remains unchanged, and an increase in supply from S to S' causes the price to decrease from 5 cents to

about 4 cents per pound, while a decrease in supply from S to S'' causes the price to increase from 5 cents to about 7¾ cents per pound. The quantity bought and sold increases with the increase in supply and decreases with the decrease in supply.

Although it is possible for demand and supply to change simultaneously in the short run, such situations are not covered specifically by the Law of Short-Run Competitive Price. But it is not difficult to figure out what would happen in such cases. Like changes in demand and supply tend to cancel each other more or less completely as far as any effect on price is

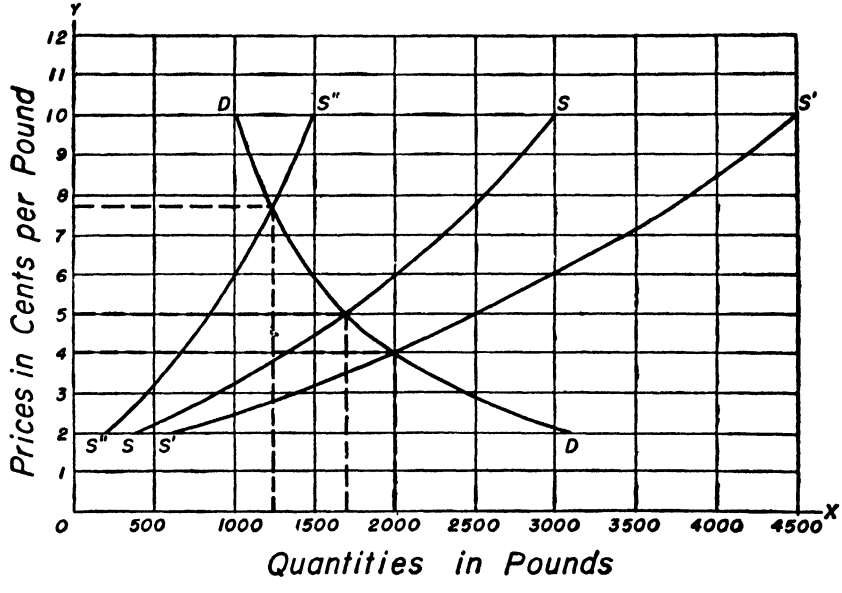

FIGURE 17.—Effect of Increased and Decreased Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions

concerned, but the quantity bought and sold changes considerably. In Figure 18 the demand for sugar increases from D to D' while supply increases from S to S'. The increase in demand tends to increase the price, while the increase in supply tends to decrease the price. In this case the two changes together leave the price unchanged at 5 cents per pound although the quantity bought and sold increases from 1,700 pounds to 2,550 pounds. In other cases of simultaneous increases in demand and supply the price might rise or fall slightly. The case of a simultaneous decrease in demand and supply is not illustrated but it may be readily understood that here also the changes in demand and supply cancel each other to a considerable extent so that price is not likely to change greatly, although the quantity bought and sold would decrease considerably.

Opposite changes in demand and supply affect the quantity bought and sold only slightly, if at all, but they exert a considerable influence on price, since both changes tend either to increase or to decrease the price. In Figure 19 the demand for sugar decreases from D to D'', while the supply of sugar increases from S to S'. The decrease in demand alone would lower the price of sugar from 5 cents to 3½ cents per pound and the increase in supply alone would lower the price to 4 cents per pound. Under the influence of the changes in both demand and supply, the price falls to slightly less than 3 cents per pound. In this case, the quantity bought and

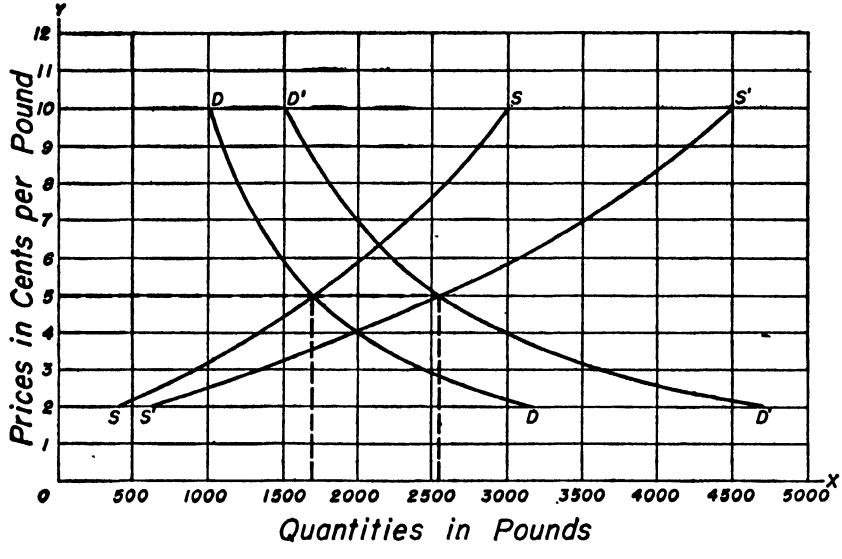

FIGURE 18.—Effect of a Simultaneous Increase in Demand and Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions

sold is decreased more by the decrease in demand than it is increased by the increase in supply, and it falls from 1,700 pounds to 1,275. In other cases, the quantity bought and sold might remain unchanged or even increase to some extent. The case of a simultaneous increase in demand and decrease in supply is not illustrated, but it is readily seen that in this situation the price of sugar would rise considerably, since each change by itself would tend to raise the price, while the quantity bought and sold would change only slightly, if at all.

The Case of Fixed Supply. There is one special case of short-run price determination under competitive conditions which merits our attention. In this case the sellers of an economic good have a fixed stock of the product on hand as usual, but the good is so perishable that the entire fixed stock

must be sold at a given time for any price obtainable. Such a perishable fixed stock is usually called a fixed supply, for the sellers would supply the same amount of the good (all they have) at any price which might

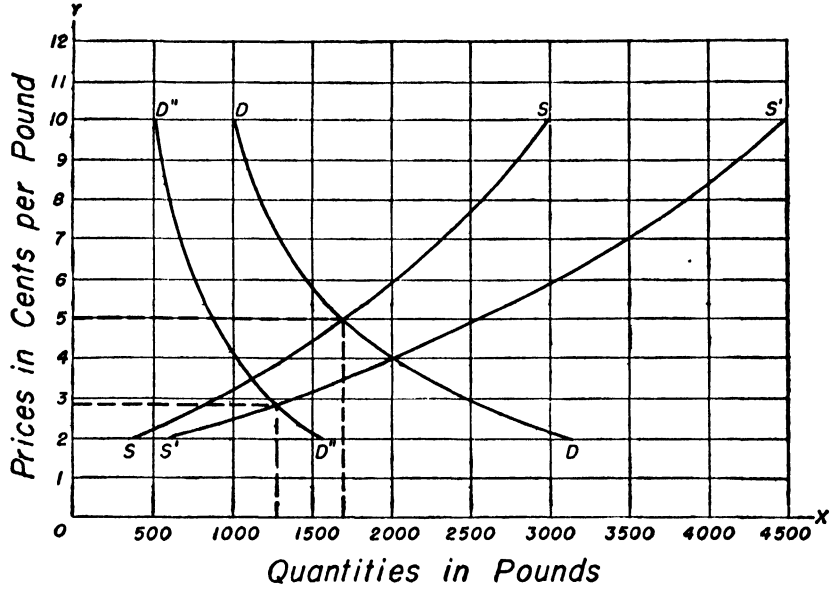

FIGURE 19.—Effect of a Simultaneous Decrease in Demand and Increase in Supply on the Price of Granulated Cane Sugar in the Market of X under Short-Run Competitive Conditions

Table 15: The Demand for Strawberries in the Market of X on a Given Day

Price Per Quart	Quantities Which Buyers Would Purchase (*in quarts*)
35¢	1,100
30¢	1,200
25¢	1,400
20¢	1,700
15¢	2,100
10¢	2,600
5¢	3,200

prevail in the market. The determination of short-run competitive price under the condition of fixed supply is rather simple. Suppose, for example, that the numerous sellers in our market of X have on their hands a fixed stock of 2,600 quarts of strawberries and that the berries are so near the

point of spoiling that they will not be salable if held until the next market day. The demand for the berries on this day is of the ordinary competitive type, as shown in Table 15. The resulting situation is diagramed in Figure 20. In this diagram, the supply curve is a straight line perpendicular to the *OX* axis at the quantity of 2,600 quarts. The price of the berries is determined at the point of intersection between the demand and supply curves, and is 10 cents per quart for 2,600 quarts. Under conditions of fixed supply, then, the only possible price for an economic good in a competitive market is the price at which the entire fixed supply will be purchased by the buyers, even though this price may be well below the cost of production of the good in question.

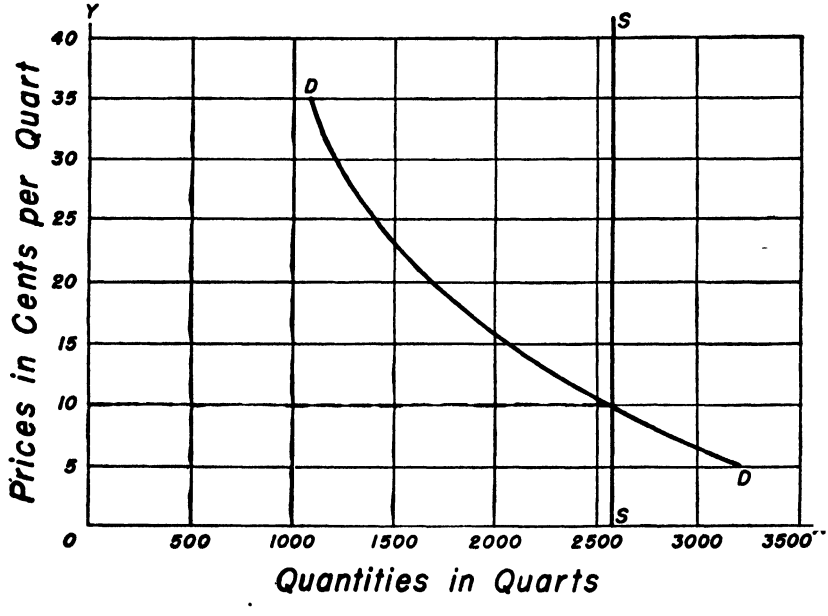

FIGURE 20.—Short-Run Competitive Price Determination, under Conditions of Fixed Supply, in the Market of X on a Given Day

Students often object to this conclusion and point out that the sellers would be much better off if, instead of selling the entire fixed supply at any price obtainable, they destroyed or threw away part of their fixed supply in order to sell the remainder at a high price. Thus 1,100 quarts sold at 35 cents per quart would bring in total receipts of $385 for the sellers, while 2,600 quarts sold at 10 cents per quart would give total receipts of only $260. There is no doubt then that sellers would gain in many cases by selling only part of a fixed supply at a high price and that they would be happy to follow this course of action if it were possible to do so. However, this procedure would be impossible under competitive conditions of the

market because competition requires that the various sellers act independently of each other. The destruction of part of a fixed supply in order to secure a high price for the remainder requires collusion among the sellers. As long as competitive conditions exist, the individual seller has no incentive to destroy part of his holdings. To do so would merely reduce the quantity which he had to sell without producing any improvement in the price per unit at which the good would sell.

PRICE DETERMINATION IN THE INTERMEDIATE PERIOD

Demand Conditions. We are now ready to discuss the determination of prices in the intermediate period under competitive conditions of the market, starting with price determination from the point of view of the individual competitive firm and moving on to price determination for a whole competitive industry composed of many firms. In passing from the short run to the intermediate period we find that the demand analysis remains virtually unchanged. We still think of demand schedules and curves as being applicable to a given market at a given time, but it is now necessary to regard the quantities which buyers will purchase, according to these schedules and curves, as amounts per month or per year rather than as amounts per day. The demand for the product of the individual competitive firm is also of the same nature in the intermediate period as in the short run. Under competitive conditions there are a great many firms producing and selling exactly the same economic good, and the demand for the product of any individual firm is a horizontal straight line parallel to the *OX* axis in a diagram and fixed at the level of the prevailing price. The individual firm can sell as much or as little as it pleases at this price, but it cannot sell any units at a higher price and it would be swamped with business at any lower price. Under this condition of perfect elasticity in the demand for the product of the individual firm, marginal revenue and average revenue are always equal, as we have seen.

Cost and Supply Conditions. In the intermediate period, either a single firm or a whole industry has a fixed amount of plant and equipment (as well as fixed amounts of certain other productive factors) and a fixed productive capacity, so that changes in output can result only from speeding up or slowing down the rate at which the fixed productive facilities are operated. In this period, each firm has some costs which remain fixed in total amount regardless of changes in output, and other costs whose total amount varies directly with output. The distinction between average and marginal costs is also important in this period since average and marginal cost will be equal only when average cost is at its minimum; i.e., when the firm is operating at normal or optimum output. The behavior of average variable, average, and marginal costs, and the relationships between these types of cost, were discussed in detail in the preceding chapter and

will not be reviewed here. We should recall, however, that the supply curve for the individual competitive firm in the intermediate period consists of that part of its marginal cost curve which lies above the curve of average variable cost, while the supply curve for a whole competitive industry is merely a summation of the supply curves of all the individual firms in the industry.

Price Determination for the Individual Firm. In Figure 21 we show the average variable, average, and marginal cost curves which might represent the different rates at which the individual competitive firm could operate with its fixed productive facilities in the intermediate period. If the perfectly elastic demand curve for the product of the individual firm at the level of the ruling price is as shown by the line AR—MR in Figure 21, this individual firm will be able to operate at normal output and its output will be *OA* units. The price per unit, *AP*, is equal to average and marginal cost per unit of product on the one hand and to average and marginal revenue per unit of product on the other. If the demand for the product of the industry should be greater, so that the perfectly elastic demand for the product of the individual firm at the level of the prevailing price is represented by the line *AR'—MR'*, the firm will be induced to achieve a rate of operation in excess of normal output. It will be willing to expand production as long as the unit price obtainable exceeds the marginal cost per unit of product. The expansion of output will stop at *OB* units because any further units of product would add more to the costs of the firm than to its revenue or income. In other words, marginal cost and marginal revenue are equal at an output of *OB* units, and the price per unit will be BP'. This price, although equal to marginal cost per unit, is well above the average cost per unit of product at the output of *OB* units, and our individual firm will be operating at a profit.

If the demand for the product of the whole industry is relatively weak, so that the perfectly elastic demand for the product of the individual firm at the level of the ruling price is represented by the line *AR''—MR''*, the individual firm finds that there is no rate of operation for its business at which average cost per unit of product will be fully covered by the obtainable price. At first thought this might suggest the desirability of closing down the productive operations of the firm altogether, but we must remember the presence and influence of fixed costs. If there is some output at which the ruling price will cover all of the variable costs of the firm and part of its fixed costs, the firm will sustain a smaller loss by producing and selling this output than by not producing at all. A glance at Figure 21 shows that there are several possible outputs at which the ruling price, represented by AR''—MR'', will more than cover average variable cost. However, as usual, the most favorable output is that at which marginal cost equals marginal revenue (and price), and the firm will produce an

output of OC units, sell at the ruling price (CP'') per unit, and thereby sustain the minimum possible loss under this condition of demand. To stop short of an output of OC units would be to fail to add to output some units of product whose marginal cost would be more than covered by the

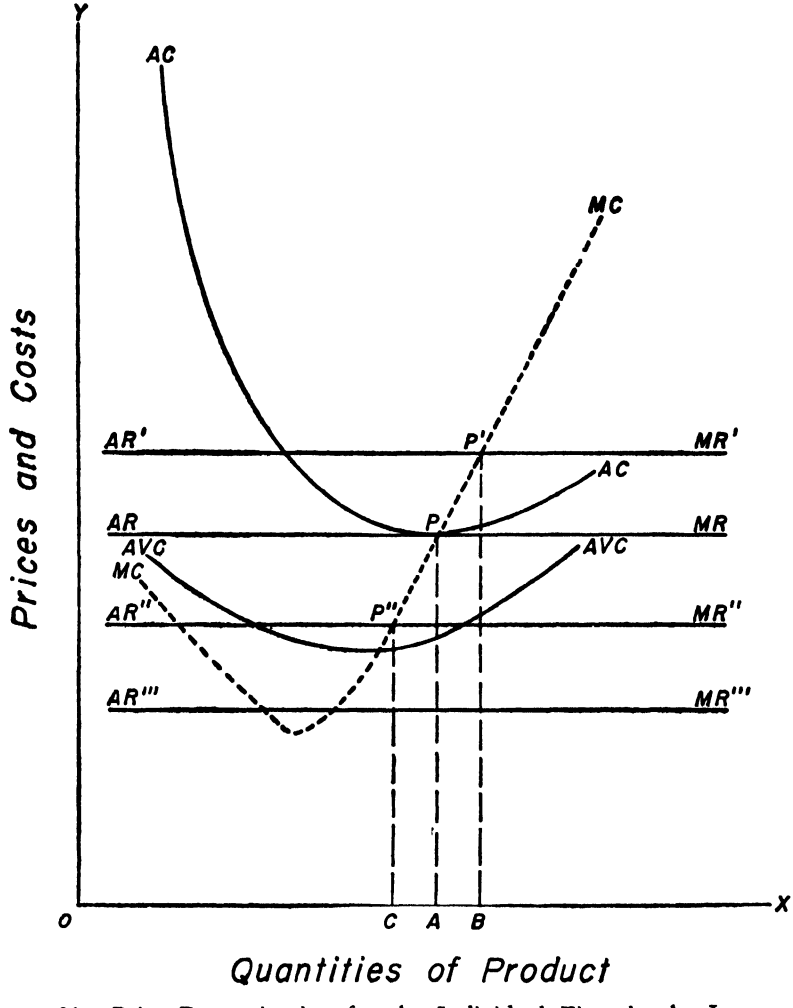

FIGURE 21.—Price Determination for the Individual Firm in the Intermediate Period under Competition

marginal revenue obtainable from their sale, while the production of any output larger than OC would involve the production of some units which would add more to the firm's costs than to its income.

Finally, if the demand for the product of the whole industry is so weak that the perfectly elastic demand for the product of the individual

firm at the level of the ruling price is represented by the line *AR'''—MR'''* in Figure 21, our individual firm would be better off not to operate at all for the time being. There are, of course, some outputs at which the ruling price would more than cover marginal cost per unit of product, but there is no output at which the price would cover average variable cost. Since the variable costs of the firm can be eliminated entirely by failing to operate, it is clear that the firm would sustain a larger loss by operating at any output than by closing down completely until demand conditions change.

Price Determination for a Whole Competitive Industry. We may now consider price determination in the intermediate period for the whole competitive industry of which our individual firm is a member. The supply curve for the whole industry is a summation of those parts of all the individual firms' marginal cost curves which lie above the level of average variable cost. As such, it may well take the form of the curve *S* in Figure 22. In this diagram we must remember that the quantity units along the *OX* axis are much greater than in Figure 21, which represented price determination for the individual competitive firm. That is, if the outputs in Figure 21 are visualized in thousands of units, those in Figure 22 must be taken to represent hundreds of thousands or millions of units. The supply curve *S* for the whole industry indicates that the quantity produced and sold would vary directly with the price in the intermediate period, thus conforming with the Law of Supply. That is, the higher the price of the product the greater will be the outputs which all the firms can produce without having their marginal costs exceed the price, and vice versa.

The demand curve for the product of the whole industry is not at all like that for the product of the individual competitive firm. The firm can sell as much or as little as it pleases at the price ruling in the market, but the whole competitive industry cannot. That is, the demand for the product as a whole shows that the quantity which buyers would purchase varies inversely with the price, in conformity with the Law of Demand. If the total demand for the product is that represented by the line *D* in Figure 22, the price of the product will be *OP* per unit and the output of the whole industry will be *OL* units. This price is the same as that represented by *AR—MR* in Figure 21, which caused our individual firm to operate at normal output. Needless to say, however, this same price might cause other individual firms to operate at a rate beyond or short of normal output, because of differences in marginal costs between firms in the intermediate period.

If the total demand for the product increased to that represented by the line *D'* in Figure 22, the price of the product would increase to *OP'* and the output of the industry would increase to *OM* units. This price is the same as that represented by *AR'—MR'* in Figure 21, which caused our individual firm to produce an output of *OB* units, or at a rate in excess

of normal output. In the case of certain high-cost firms, the high price of *OP'* may bring about an increase in production only to or toward normal output, but it will cause most individual firms to expand production to a point beyond normal output. Operating at a rate beyond normal output and receiving a price which covers marginal cost, these firms will find that they are receiving a price which more than covers average cost per

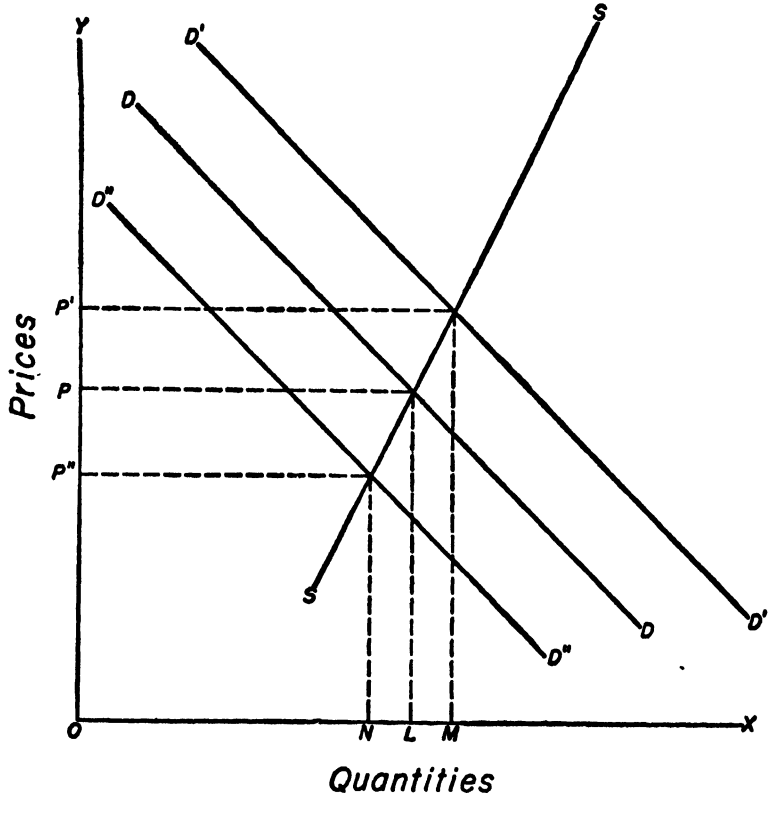

FIGURE 22.—Price Determination for a Whole Industry in the Intermediate Period under Competition

unit of product and that they are making profits. However, the firms in the industry have no incentive to produce more than a total of *OM* units, since additional units would command a price insufficient to cover marginal cost per unit, and the intermediate period is too short to permit the productive capacity of the whole industry to be expanded by means of the entrance of new firms. Hence the price of *OP'* is an equilibrium price for the whole competitive industry in the intermediate period under the demand represented by *D'* even though this price leads to profits for most

of the individual firms in the industry. In fact, it is this widespread existence of profits which is relied on to cause the productive capacity of the industry to increase *in the long-run period* if the increase in demand is relatively permanent.

If the total demand for the product decreased to that represented by the line D'' in Figure 22, the price of the product would decrease to OP'' and the output of the industry would decrease to ON units. This price is the same as that represented by AR''—MR'' in Figure 21, which caused our individual firm to produce an output of OC units, or at a rate short of normal output. In the case of certain low-cost firms, the low price of OP'' may still permit operation at or beyond normal output, but it will cause most individual firms to contract production to a point short of normal output. Operating at a rate short of normal output and receiving a price which covers marginal cost, these firms will find that the price fails to cover average cost per unit of product so that they are operating at a loss. However, under the assumed conditions, the firms are better off when they produce such an output than they would be if they stopped operating altogether, and the intermediate period is too short to permit the productive capacity of the industry to be decreased by means of the exit of firms. Hence the price of OP'' is an equilibrium price for the industry in the intermediate period under the demand represented by D'' even though this price causes losses for most of the individual firms in the industry. In fact, it is this widespread existence of losses which is relied on to cause the productive capacity of the industry to decrease *in the long-run period* if the decrease in demand is relatively permanent.

QUESTIONS AND PROBLEMS

1. "In the short run under competition the equilibrium price for an economic good is that price at which market demand and market supply are equal." Explain.

2. Assume that the following demand and supply schedules represent short-run competitive conditions in the market for a certain economic good:

Price Per Unit	Quantities Which Buyers Would Purchase (*in pounds*)	Quantities Which Sellers Would Offer (*in pounds*)
50¢	6,500	16,000
45¢	7,000	14,000
40¢	7,500	12,000
35¢	8,000	10,500
30¢	9,000	9,000
25¢	10,500	7,500
20¢	12,000	6,000
15¢	14,000	4,500

(a) What will be the price of the good under the assumed conditions? Why?
(b) Why could the price not be higher or lower?
(c) What would the price be if the demand for the good increased by 50 per cent?

3. "The Law of One Price is the worst kind of 'theoretical' law for it is obvious that it doesn't work in practice." Discuss.

4. Illustrate the Law of Short-Run Competitive Price with appropriate diagrams.

5. "A simultaneous increase in demand and decrease in supply is likely to leave the price of an economic good virtually unchanged in the short run, but the quantity bought and sold will change considerably." Show whether you agree.

6. "Under conditions of fixed supply it is unreasonable to assume that competing sellers will sell their entire stock of a good at whatever price can be obtained, for they would be much better off in many cases to destroy or throw away a part of their stock in order to sell the remainder at an enhanced price." Do you agree? Explain.

7. Explain the differences which exist between the diagram representing price determination for the individual firm and that for price determination for a whole industry in the intermediate period under competitive conditions.

8. "If it is worth while for him to produce at all, the competitive enterpriser will always be willing to expand output to the level at which price equals marginal cost in the intermediate period." Show whether you agree.

9. "In the intermediate period under competition it may be better for an enterpriser to produce and sell his good at a loss than to close down altogether." Explain.

10. "An enterpriser whose total costs are made up largely of fixed costs will be less willing to sell his product at a price below average cost of production in the intermediate period than an enterpriser whose total costs are made up largely of variable costs." Do you agree? Explain.

11. A certain radio manufacturer sells a given model in this country for $100 and markets the same model in other countries for $75, which is less than average cost of production per unit. Why may the manufacturer be anxious to make foreign sales at this price? Explain.

12. "If the output at which price is equal to his marginal cost is in excess of normal output, the competitive enterpriser will be operating at a profit in the intermediate period." Explain.

13. "The competitive enterpriser in the intermediate period cannot afford to produce an output at a price equal to marginal cost if that price is less than average variable cost." Show whether you agree.

14. "A given price for an economic good may be regarded as an equilibrium price in the intermediate period under competition even though it causes widespread profits or losses for the individual firms in the industry." Explain.

See References for Further Reading at the end of Chapter XVI.

XIV

The Determination of Prices under Competition (*Continued*)

PRICE DETERMINATION IN THE LONG RUN

We come now to the study of price determination in the long run under conditions of competition. The long run, as we have seen, is that period of time which permits the most complete adjustment of the production of an economic good to the demand for it. In the intermediate period, production may be adjusted to demand only by operating faster or slower with fixed quantities of certain productive facilities in an industry, but in the long run the amounts of all productive agents in an industry, and hence its productive capacity, may be changed. The long run, in other words, is long enough for new firms and agents of production to come into an industry or for old firms and agents of production to pass out of an industry, according to whether an expansion or a contraction of its productive capacity is necessary.

Price and Cost in the Long Run. In the intermediate period, as we have noted, the price of an economic good may be regarded as an equilibrium price under competition even though it is above or below the average cost of production of most of the individual firms in the industry so that profits or losses are widespread. However, in the long run the equilibrium price of an economic good must be equal to both average and marginal costs of all the firms in a competitive industry. If the price of the good is temporarily above the average cost of most of the firms in a competitive industry (though equal to marginal cost for the outputs which they produce), profits are made by these firms, and the agents of production in use earn more than those which are employed in other fields. As a result, new firms and agents of production will be attracted to the industry, its size, productive capacity, and output will be increased, and the price will fall back to a cost of production level once more. If the price of the good remains below the average cost of production of most of the firms in a competitive industry for any considerable time (though equal to marginal costs for the outputs which they produce), losses are sustained by

these firms and the productive agents in use earn less than those which are employed in other fields. As a result, some firms and productive agents will leave the industry, its size, productive capacity, and output will be decreased, and the price will rise to a cost of production level once more.

On the basis of these considerations, we may state the Law of Long-Run Competitive Price, which holds that *the price of an economic good in the long run under competition, as determined by demand and supply, tends to equal cost of production per unit for all the firms in the industry.* Under any condition of demand, each firm in a competitive industry in the long run tends to achieve an optimum size (though not exactly the same size from firm to firm) and optimum output. Regardless of differences in optimum size from firm to firm and in the slope of cost curves, all firms in a competitive industry, as we have seen, tend to have the same minimum long-run average cost, so that it is quite possible for price to equal long-run average cost for all the firms in the long run. Moreover, when the long-run average cost of a firm is at its minimum, it is equal to long-run marginal cost, to intermediate-period average cost for the optimum size and output of the firm, and to intermediate-period marginal cost for the optimum size and output of the firm. If price is equal to the long-run average cost of the firm, it is equal to these other costs as well. Hence in stating the Law of Long-Run Competitive Price it is not necessary to specify any one type of cost of production.

The student should be careful to understand exactly as it is stated the proposition with regard to the relationship between cost of production and price in the long run under competitive conditions. It has not been said, for example, that the price of an economic good is *determined* by cost of production in the long run under competition. Cost of production, operating through supply, is emphasized in the discussion of long-run price determination under competition, but this does not mean that cost of production determines price. Demand is essential to the determination of a long-run price equal to cost of production. If an enterpriser produced, at a cost of one thousand dollars per unit, a machine to amuse children by blowing soap bubbles, it would not necessarily follow that these machines could be sold for a price which would cover cost. Moreover, while changes in cost of production are certain to affect price in the long run under competition, it should be remembered that changes in industrial size, productive capacity, and cost of production usually follow upon changes in demand conditions for an economic good. The proper statement is, then, that the price of an economic good in the long run under competition is determined by demand and supply in such a way that it tends to equal the cost of production per unit of the firms in the industry.

It should be clear from the discussion of competitive price determination in the short-run and intermediate period that the Law of Long-Run

Competitive Price is merely a statement of tendency and does not mean that price is actually always equal to cost of production under competitive conditions. Cost of production is merely the normal level or position around which the actual price of an economic good fluctuates. Somewhat analogous is the normal position of the pendulum on a grandfather's clock, which is straight up and down, but when the clock is running, the pendulum spends very little time in this position. Most of the time it is going away from or coming back to this position. The price of an economic good under competitive conditions may also spend very little time at the cost of production level. Of course the oscillations of price around cost of production are not rhythmic and regular as are the movements of a pendulum, and price may spend more time on one side of the cost of production level than on the other.

After comparing the movements of price and production around their equilibrium positions with the swinging of a stone hanging freely by a string, one noted economist said: "But in real life, such oscillations are seldom as rhythmical as those of a stone hanging freely by a string; the comparison would be more exact if the string were supposed to hang in the troubled waters of a mill-race, whose stream was at one time allowed to flow freely, and at another partially cut off. Nor are these complexities sufficient to illustrate all the disturbances with which the economist and the merchant alike are forced to concern themselves. If the person holding the string swings his hand with movements partly rhythmical and partly arbitrary, the illustration will not outrun the difficulties of some very real and practical problems of value." [1] However, cost of production is the equilibrium position for the price of an economic good in the long run under competition in the sense that, if the price is displaced from this position by very far for very long, forces tend to be set in motion more or less automatically which will sooner or later cause the price to move back toward the position of equilibrium.

Cost of Production and Size of Industry. Since the price of an economic good in the long run under competition tends to equal cost of production per unit for all the firms in the industry, and since this tendency depends upon the expansion or contraction of the size and productive capacity of the industry in response to prices higher or lower than cost of production, the question of what happens to cost of production for all the firms as an industry expands by means of the entrance of new firms or contracts by means of the exit of old firms is very important for price determination in the long run under competition. As we noted in Chapter XII, either the expansion of an industry by means of the entrance of firms or the contraction of an industry by means of the exit of firms involves three possibilities with respect to the cost of production of all the firms. That is,

[1] Alfred Marshall, *Principles of Economics*, p. 346.

this cost of production may increase, decrease, or remain constant. These long-run cost tendencies are named according to what happens to the cost of production of all the firms as the industries *increase* their size and productive capacity.

Thus an industry is said to operate under conditions of increasing cost in the long run if an expansion in its size and productive capacity by means of the entrance of firms results in an increase in the cost of production per unit of product for all the firms.[2] If the industry contracted in size and productive capacity by means of the exit of firms, cost of production per unit of product would decrease for all the firms. If an industry operated under conditions of decreasing cost in the long run, an expansion of its size and productive capacity by means of the entrance of firms would lead to a fall in the cost of production per unit of product for all the firms, while a contraction of size and productive capacity by means of the exit of firms would produce an increase in the cost of production per unit of product for all the firms. Finally, if an industry operated under conditions of constant cost in the long run, an expansion or contraction in its size and productive capacity by means of the entrance or exit of firms would leave cost of production per unit of product unchanged for all the firms.

Since a good deal of confusion exists among students of economics (and even among economists) as to the exact meaning of these three long-run cost tendencies, we shall pause here to explain what these tendencies do not mean. In the first place, these cost tendencies must not be confused with what happens to cost of production per unit in actual historical periods of time. It may cost less, on the average, to raise a bushel of wheat today than it did in 1920, but this does not prove, by any means, that wheat raising is an industry which operates under conditions of decreasing cost. The general price level may be lower today and the purchasing power of money higher than was the case in 1920, while, in the intervening period, important technological changes may have occurred which have lowered the cost per bushel of raising wheat. In our value analysis we assume that the general price level and purchasing power of money remain constant, and we exclude any important new inventions or changes in methods of production from our analysis of long-run cost tendencies, for such things may lower cost of production per unit in any industry, regardless of type. We are interested only in what happens to cost of production as an industry changes in size and productive capacity in the long run under a constant price level and under a given technology of production.

Again, the three long-run cost tendencies must not be confused with what happens to average cost of production per unit when a firm with a

[2] Cost of production may be interpreted here as long-run minimum average cost, but we have seen that this cost, under equilibrium conditions, is equal to several other types of cost.

fixed amount of plant and equipment, and with certain fixed costs, changes its rate of operation from part capacity to optimum output or to a point beyond optimum output. Any firm with certain fixed productive facilities, and with certain fixed costs resulting therefrom, will experience decreasing costs per unit of product as it increases output from part capacity to optimum output, and it will experience increasing costs per unit of product as it increases production beyond the optimum output which is associated with its size at the time. These events have to do with the costs of production of the individual firm in the intermediate period. They occur for the individual firm in this period regardless of the long-run cost tendency under which the industry as a whole may operate, and have nothing to do with long-run conditions of increasing, decreasing, or constant cost. It is therefore completely wrong to say that an industry of decreasing costs is one which is operating at some point in the declining portion of its intermediate-period average cost curve, or that an industry of increasing costs is one which is operating somewhere in the ascending portion of its intermediate-period average cost curve.[3]

Finally, the three long-run tendencies must not be confused with what may happen to the cost per unit of product as an individual firm starts out in an industry, develops toward maturity, and eventually fades away. It is probably true that many firms operate at lower cost after they have been in business for some years than when they first start out, and that aged firms sometimes show a decline in efficiency and an increase in cost per unit of product, but these things do not concern us here. We are interested in the costs of the individual firms only as these costs are affected by the expansion or contraction of the whole industry under competition by means of the entrance or exit of firms.

Price Determination in the Long Run under Constant Cost. An industry is said to operate under conditions of constant cost in the long run only if the costs of the individual firms are not affected by changes in the

[3] However, this mistake is frequently made. In many books the stock example of an industry which operates under conditions of decreasing cost is the railroad industry. Each railroad, we are told, has a heavy investment in tracks, rolling stock, terminal facilities, and other things, and has large fixed costs as a result. If it operates at only part capacity, the cost per unit of service is high, since the large fixed costs must be spread over a small number of units of goods and passengers transported. If the railroad can increase its business so as to operate at full capacity, the fixed costs are spread over a larger number of units transported and the cost per unit of service declines. Thus, it is argued, the railroad industry is one of decreasing cost. We must disagree emphatically with this conclusion. There is no doubt that the cost per unit of service will be less when a railroad operates at full capacity than when it runs at part capacity, but this proves nothing as to the long-run nature of the industry since the same thing will tend to be true of any industry whose firms have fixed costs in the intermediate period. What we are interested in, from the long-run point of view, is what happens to cost of production per unit of service as the railroad industry as a whole increases or decreases its size and productive capacity, and not what happens to cost per unit as individual railroads change their rate of operation under a given size and productive capacity. Now the railroad industry might actually be one of decreasing cost in the long-run sense, too, but this is not shown by the kind of illustration noted above.

number of firms in the industry. In other words, firms can enter such an industry without bidding up the cost of any productive agents used in the industry and firms can leave the industry without causing a decrease in the cost of any of these productive agents. Such results are possible only if the industry as a whole is so small that it uses only a minute fraction of the total available quantities of certain productive agents. While there are not likely to be many industries which operate under conditions of constant cost in the long run, we may take the production of handmade suits

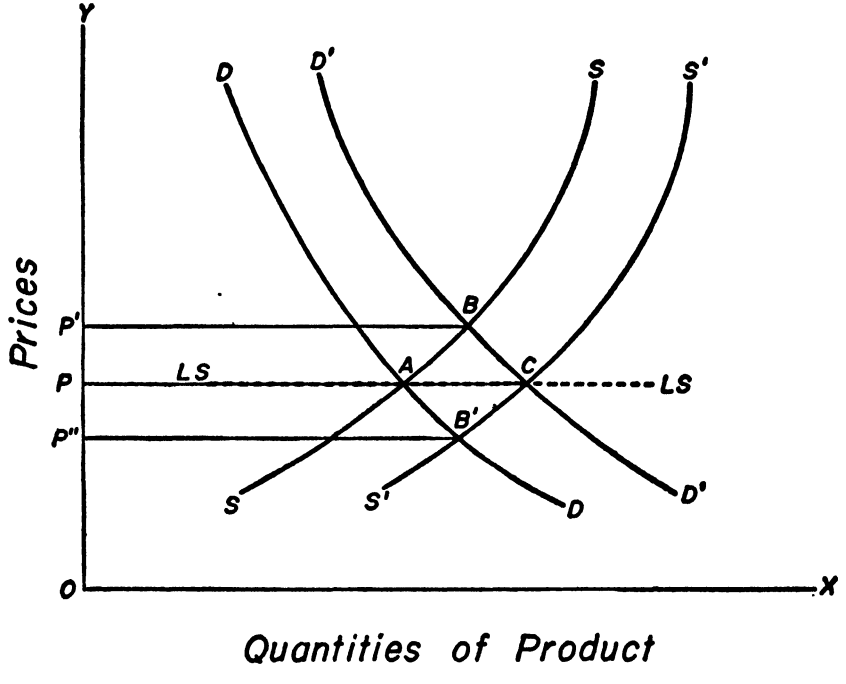

FIGURE 23.—Long-Run Price Determination in a Competitive Industry under Conditions of Constant Cost

of clothes for men as a tentative example. The quantities of needles, pins, thread, cloth, labor, and land space used by this industry are small relative to the total quantities of these things (of appropriate grades) which are available in the whole economy, and it may be that the entrance or exit of firms would not affect the cost of these productive agents to the individual firms.

In Figure 23, we start with a situation in which the industry is in a state of intermediate-period and long-run equilibrium with a demand represented by *D* and a supply represented by *S*. The price is *OP* and is equal to long-run minimum average cost for the individual firms. The output

of the industry is PA units. Now suppose that the demand increases to D' and the price of the product rises to OP'. In the intermediate period the individual firms in the industry, operating on the principle of producing as many units as possible without having marginal cost exceed price, expand their outputs somewhat, and the output of the industry increases from PA to $P'B$. However, in producing beyond optimum output at a price equal to marginal cost, the individual firms find that their average cost is more than completely covered by the price, and considerable profits are made. This situation can continue to prevail in the intermediate period, but in the long run new firms are attracted to the industry. This causes the supply curve to move to the right (indicating an increase in supply), for it is now the summation of the marginal cost curves of a larger number of firms than formerly. With the increase in supply from S to S', the output of the industry becomes PC units and the price falls back to OP per unit. This price is equal to the long-run minimum average cost of the individual firms, profits are no longer made, and the industry is once more in both intermediate-period and long-run equilibrium.

We may also use Figure 23 to illustrate the contraction of the size and productive capacity of the industry. For this purpose, we take D' and S' to be the original demand and supply curves representing intermediate-period and long-run equilibrium. The price of the product is OP, and the output of the industry is PC units. Now suppose that the demand for the product decreases from D' to D and the price falls to OP''. The individual firms will reduce their outputs in the intermediate period, but they will still produce the outputs at which marginal cost is equal to the price, even though such outputs must be produced at a loss.[4] Hence the output of the industry will be $P''B'$. In the long run, however, some firms and agents of production will leave the industry. This will cause the supply curve to move to the left (indicating a decrease in supply), for it is now the summation of the marginal cost curves of a smaller number of firms than formerly. With the decrease in supply from S' to S, the output of the industry becomes PA units and the price goes back up to OP per unit. This price is equal to the long-run minimum average cost of the individual firms, losses are no longer sustained, and the industry is once more in both intermediate-period and long-run equilibrium. Since either an expansion or a contraction of the industry leaves the minimum average cost of the firms and the price of the product unchanged at OP *in the long run*, the long-run cost and supply curve of the industry is a series of points such as A and C in Figure 23 and is shown by the broken line LS.

Price Determination in the Long Run under Increasing Cost. An in-

[4] Since S' is the summation of only those parts of the individual firms' marginal cost curves which lie above the level of average variable costs, the firms will lose less by producing the output at which marginal cost is equal to price than by closing down.

dustry is said to operate under conditions of increasing cost in the long run if the costs of all the individual firms tend to increase as the industry expands in size and productive capacity by means of the entrance of new firms. The expansion of the industry and the entrance of firms require the use of larger quantities of productive agents than formerly in the industry as a whole. The increase in cost for the individual firms may result from bringing into the industry quantities of productive agents which are not as well suited for the purposes of the industry as those already in use, or it may result from the acquisition of high-grade productive agents through the process of bidding them away from other industries. In the first case, cost increases because of the low quality and productivity of the newly acquired agents of production. In the second case, the attempt to acquire additional quantities of high-grade agents already in use in other industries results in an increase in the total demand for these agents and an increase in their price or cost in all their alternative uses. In either case, the high cost (per unit of product) of the additional agents is reflected in the costs of all the firms, since all firms in a competitive industry tend to have the same minimum average cost in the long run. On the other hand, the contraction of the size and productive capacity of the industry and the exit of firms would be expected to lower the costs of all the firms, either because poorer grades of certain productive agents would be dropped from use or because the contraction of the industry, other things equal, would cause some decrease in the total demand for certain high-grade productive agents and a decrease in their price or cost in all their alternative uses. Most industries are likely to operate under conditions of increasing cost in the long run.

Figure 24 illustrates the determination of prices in the long run under conditions of increasing cost. We start with the industry in a condition of intermediate-period and long-run equilibrium with a demand represented by D and a supply represented by S. The price is OP and is equal to long-run minimum average cost for the individual firms. The output of the industry is PA units. Now suppose that the demand increases to D' and the price of the product rises to OP'. The individual firms already in the industry expand their outputs as much as they can without having their marginal cost exceed this price, considerable profits are made, and the output of the industry expands to $P'B$ units. In the long run, the existence of profits attracts additional firms to the industry, their entry requires a shift of productive agents to the industry, and the cost of certain productive agents is raised for all the firms in the industry. Each firm already in the industry finds that its cost curves are on a higher level than formerly.

In this situation, two influences are at work on the supply curve S. The raising of the marginal cost curves of the individual firms originally in the industry tends to move S to the left, but this influence is more than

completely offset by adding in the outputs which would be produced at various prices by the new firms which have just entered the industry. The net result is that the supply curve shifts to the right, indicating an increase in supply, and takes the position of the curve S' in Figure 24. With this increase in supply, the output of the industry becomes $P''C$ units in the long run and the price of the product falls to OP'' per unit. This price is equal to the long-run minimum average cost of all the individual firms, profits are no longer made, and the industry is once more in both inter-

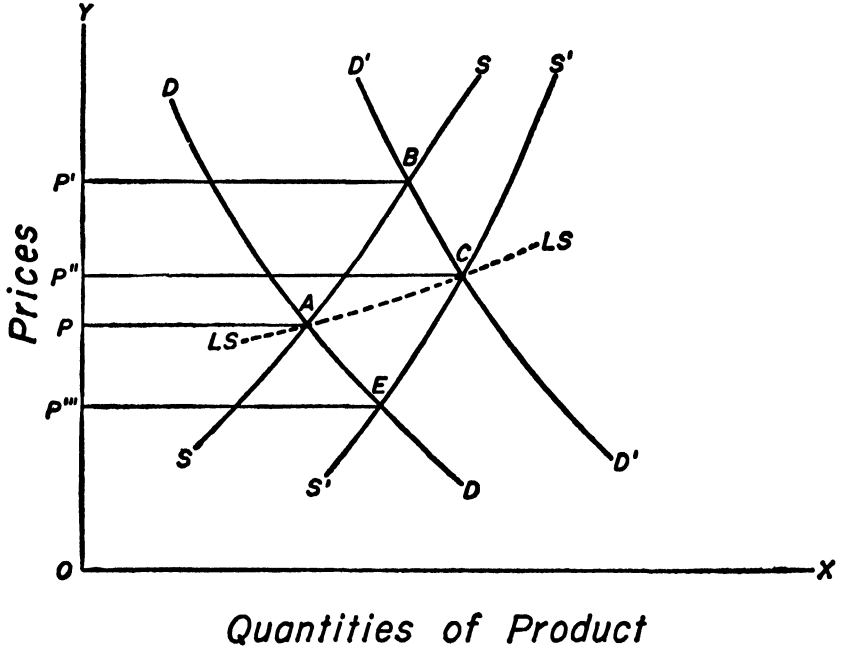

FIGURE 24.—Long-Run Price Determination in a Competitive Industry under Conditions of Increasing Cost

mediate-period and long-run equilibrium. However, the minimum average cost and price per unit is higher than it was when the industry as a whole had a smaller size and productive capacity.

In analyzing the effects of a decrease in the size and productive capacity of the industry, we take D' and S' in Figure 24 to be the original demand and supply curves representing intermediate-period and long-run equilibrium. The price of the product is OP'' and the output of the industry is $P''C$ units. Then the demand for the product decreases from D' to D and the price falls to OP'''. The individual firms reduce their outputs in the intermediate period but they still produce the outputs at which marginal cost is equal to the price, even though such outputs must be produced at

a loss. The output of the industry is temporarily $P'''E$. In the long run, some firms and agents of production will leave the industry, and the cost of certain productive agents will be less than formerly for all firms in the industry. Each firm's cost curves will be on a lower level than formerly. On net balance, the supply curve will shift to the left (since the departure of firms and loss of their outputs will more than cancel the effect of lowering the marginal cost curves of the remaining firms) and takes the position of S in the diagram. With this decrease in supply, the output of the industry becomes PA units, and the price per unit of product goes up to OP. This price is equal to the long-run minimum average cost of the individual firms, losses are no longer sustained, and the industry is once more in both intermediate-period and long-run equilibrium. However, the minimum average cost and price per unit is lower than it was when the industry as a whole had a larger size and productive capacity. Since the minimum average cost of the firms in such an industry and the price of the product tend to increase as the industry expands in size and productive capacity and to decrease as the industry contracts, the long-run cost and supply curve is a series of points, such as A and C in Figure 24, and is shown by the broken line LS.

Price Determination in the Long Run under Decreasing Cost. An industry is said to operate under conditions of decreasing cost in the long run if the costs of all the individual firms tend to decrease as the industry expands in size and productive capacity by means of the entrance of new firms. The expansion of the industry and entrance of firms requires, as usual, the use of larger quantities of productive agents than formerly in the industry as a whole, but in this case the larger quantities of productive agents can be obtained at lower cost per unit than the smaller quantities formerly used. As we have seen previously, this might happen if an industry obtained certain productive agents or services from other industries which also operated under conditions of decreasing cost, or if it acquired some productive agent from a monopoly which was operating under such conditions that an increase in its output would permit a lowering of price. However cases of these kinds, if they exist at all in practice, must be quite rare.

Figure 25 illustrates the determination of prices in the long run under conditions of decreasing cost. We start with the industry in a condition of intermediate-period and long-run equilibrium with a demand represented by D and a supply represented by S. The price is OP and is equal to long-run minimum average cost for the individual firms. The output of the industry is PA units. Now suppose that the demand increases to D' and the price of the product rises to OP'. The individual firms already in the industry expand their outputs as much as they can without having their marginal cost exceed this price, considerable profits are made, and the out-

put of the industry expands to $P'B$ units. In the long run, the existence of profits attracts additional firms to the industry, their entry requires an increase in the quantities of productive agents used in the industry, and the cost of certain productive agents is lowered for all the firms in the industry. Each firm already in the industry finds that its cost curves are on a lower level than formerly. Both the lowering of the marginal costs of the original firms and the entrance of the new firms into the industry tend to cause the supply curve to move to the right, indicating an increase in supply. As the supply curve increases from S to S', the output of the

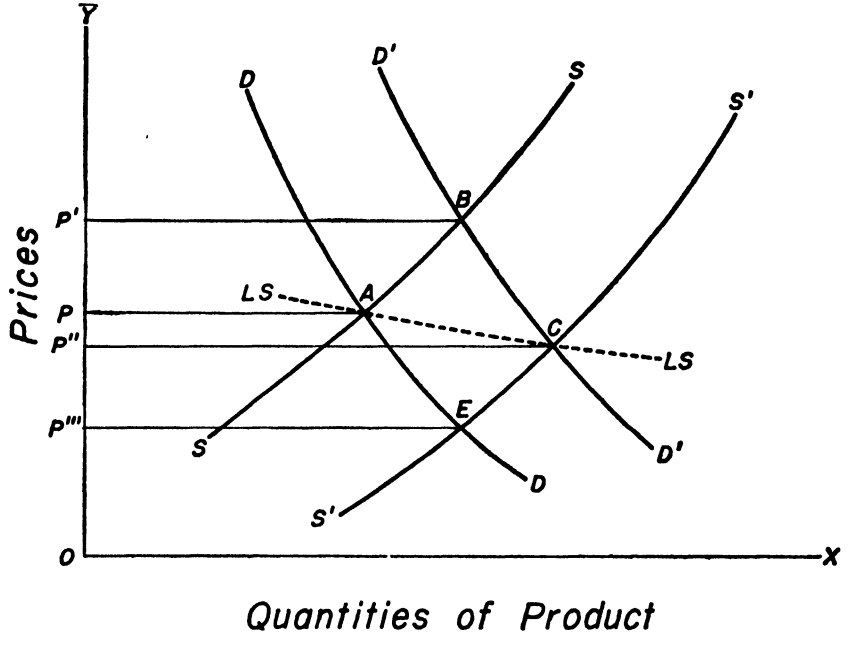

FIGURE 25.—Long-Run Price Determination in a Competitive Industry under Conditions of Decreasing Cost

industry becomes $P''C$ units in the long run and the price of the product falls to OP'' per unit. This price is equal to the long-run minimum average cost of all the individual firms, profits are no longer made, and the industry is once more in both intermediate-period and long-run equilibrium. However, the minimum average cost and price per unit is lower than it was when the industry as a whole had a smaller size and productive capacity.

In observing the results of a decrease in the size and productive capacity of the industry, we take D' and S' in Figure 25 to be the original demand and supply curves representing intermediate-period and long-run equilibrium. The price of the product is OP'' and the output of the industry is $P''C$ units. Then the demand for the product decreases from D' to D

and the price falls to OP'''. The individual firms reduce their outputs in the intermediate period, but they still produce the outputs at which marginal cost is equal to the price, even though such outputs must be produced at a loss. The output of the industry is temporarily $P'''E$. In the long run, some firms and agents of production will leave the industry, and the cost of certain productive agents will be higher than formerly for all firms in the industry. Each firm's cost curves will be on a higher level than before. The supply curve will shift to the left, both because of the higher marginal cost curves of the surviving firms and because it will be the summation of the marginal cost curves of fewer firms than formerly, and takes the position of S in Figure 25. With this decrease in supply, the output of the industry becomes PA units and the price of the product goes up to OP. This price is equal to the long-run minimum average cost of the individual firms, losses are no longer sustained, and the industry is once more in both intermediate-period and long-run equilibrium. However, the minimum average cost and price per unit is higher than it was when the industry as a whole had a larger size and productive capacity. Since the minimum average cost of the firms in such an industry and the price of the product tend to decrease as the industry expands in size and productive capacity and to increase as the industry contracts, the long-run cost and supply curve is a series of points such as A and C in Figure 25, and is shown by the broken line LS.

Changes in Long-Run Cost and Supply Curves. We may now see how important new inventions and changes in industrial methods and processes (as well as other matters) are related to the various long-run cost tendencies. The effect of such technological changes is on the *level* of the entire long-run cost and supply curves and not on their *slope*. In Figure 26, let D be the long-run demand for wheat and LS be the long-run cost and supply curve in a given state of the arts of production. The industry is taken to be one which operates under conditions of increasing cost in the long run. Now suppose that an important invention occurs which greatly facilitates wheat raising. The result is an entire new long-run cost and supply curve (LS') which is on a lower level than the former curve and indicates that the long-run minimum average cost of producing wheat will be lower than before, whatever may be the size and productive capacity of the industry as a whole. However, the new long-run cost and supply curve also indicates that the industry is still one which operates under conditions of increasing cost and that the long-run minimum average cost of producing wheat will be greater when the industry is large than when it is small. Inventions and other technological changes can lower the entire level of cost in the long run but they cannot, as some people mistakenly think, change an industry of increasing cost to one of decreasing cost.

A similar lowering of the entire long-run cost and supply curve for an industry might be produced by improvements in methods of production which did not involve new inventions or machines, or by a bounty granted to enterprisers by the government. Such a bounty in the wheat-raising industry, for example, would give wheat raisers a certain grant of money for each bushel of wheat produced, would defray a part of the cost of production per bushel, and would lower the long-run minimum average cost which the growers would have to derive from the selling price of the product. Other factors might raise the entire level of an industry's long-run

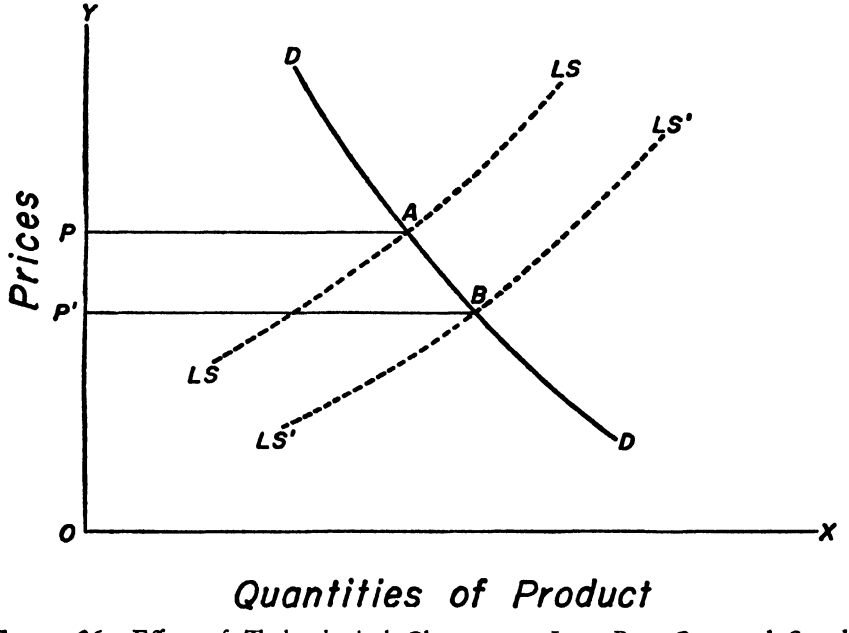

FIGURE 26.—Effect of Technological Changes on Long-Run Cost and Supply Curves under Competition

cost and supply curve without changing in any way the long-run cost tendency under which the industry operated. For example, if the government imposed a certain tax per unit of product, the long-run cost and supply curve would move to a higher level and would indicate that the minimum average cost of producing the product would be greater than before, regardless of the size and productive capacity which the industry might achieve. However, if the industry operated under conditions of increasing cost before the tax was imposed, it would continue to do so after the imposition of the tax.

Summary. In summarizing price determination under competition, we may recall that, in the short run, the price of an economic good tends to

be at a level which brings into balance the quantity which buyers are willing to purchase and the quantity which sellers are willing to dispose of out of the fixed total quantity on hand and ready for sale. The price which prevails for the good may be higher or lower than the cost at which it was produced and it may also be higher or lower than the price which will prevail when the firms have had a chance to adjust their outputs to existing conditions of demand. In the intermediate period, the individual firms have the ability to change their rates of operation and outputs, but the size and productive capacity of the firms and the industry cannot be changed. In this period the individual firms will adjust their outputs until the price is equal to their marginal costs, unless by chance the price is so low that it is less costly for them to close down altogether for the time being. From the point of view of the whole competitive industry, the price will be one which equilibrates the quantity which buyers are willing to purchase with the total output which the firms are willing to produce on the basis of their marginal costs.

However, the equilibrium price in the intermediate period may be either higher or lower than the equilibrium price which will prevail in the long run when the size and productive capacity of the industry as a whole can be adjusted to existing demand conditions. If the equilibrium price in the intermediate period, though equal to the marginal costs of the individual firms, is higher than the average costs of these firms so that profits are widespread, the size and productive capacity of the competitive industry as a whole will be expanded in the long run by means of the entrance of new firms. Conversely, if the equilibrium price in the intermediate period, though equal to the marginal costs of the individual firms, is lower than the average costs of these firms so that losses are widespread, the size and productive capacity of the competitive industry as a whole will be contracted in the long run by means of the exit of firms. In the long run, the size, productive capacity, and output of a competitive industry will tend to be such that the equilibrium price will be equal to the long-run minimum average cost of all the firms in the industry, when each firm is at the optimum size on the basis of the particular agents of production at its disposal so that all firms will have the same long-run minimum average cost of production. Finally, competitive industries are said to operate under conditions of increasing, decreasing, and constant cost in the long run according to the effect which changes in the size and productive capacities of the industries, by means of the entrance or exit of firms, will have on the long-run minimum average cost of all the firms.

INTERDEPENDENT PRICES

Joint Supply (or Joint Costs). We now turn to a brief consideration of some special cases in the field of price determination. One such case is

commonly known as joint supply or joint costs. Conditions of joint costs are said to prevail when two or more economic goods are produced with a single and indivisible cost of production. That is, certain processes of production are carried on, two or more salable economic goods result, there is a single cost of producing these products, and there is no way at all of determining what part of the total cost goes into each product. When farmers raise cotton, they produce both cotton fiber and cottonseed. There is a separate demand for each of these goods, and each commands a price. It is impossible to produce one of these goods without getting the other, and the farmer, though he knows what it costs to produce both of the goods, has no idea what it costs to produce either one. Similarly, when steers are raised, both hides and carcasses are eventually produced, the raising of sheep furnishes both wool and mutton, and so on. Articles produced under conditions of joint costs are usually called joint products.

The term "joint products" is often loosely used to include by-products, but this seems to be an improper use of the former term. Gasoline, kerosene, and lubricating oils are not really joint products, nor are beefsteaks, bone buttons, glue, gelatin and fertilizer. In the refining of petroleum, gasoline is produced first, but kerosene and lubricating oils do not automatically result from the process. Instead, further processing is required to obtain these later products and special costs are involved. In similar fashion, the processes of slaughtering and meat packing do not automatically produce the various by-products which have been listed above. Certain waste parts of the slaughtered animals are converted into a variety of other products, but extra processes of production and extra costs are required to get the by-products. Where special costs must be incurred to get some of the products, these by-products must, in the long run, sell for at least enough to cover these costs of special processing, and the various goods are not produced under conditions of joint costs in the strict sense of the term.

Now, of course, it may be objected that both cotton and cottonseed, for example, require further and special processing before they are ready for final consumption. This is true, but the point is that, since both the cotton and the cottonseed are salable economic goods in their raw state before there are any special costs of either one, they are therefore true joint products at that point. The cottonseed must be separated from the cotton, or vice versa, but the cost of that process cannot logically be charged entirely to either of the products. In the meat-packing business, it can be said that the meat and the waste materials out of which by-products are made are joint products if these materials can be sold without further processing, but it is not true that the meat and the finished by-products are really joint products. Most of the examples of articles produced under genuine conditions of joint costs are found in agriculture and are likely to be produced under conditions of increasing cost in the long run.

Price Determination under Conditions of Joint Cost. In the long run the price of an economic good is supposed to equal the long-run minimum average cost of the firms in the industry under competitive conditions, but joint products, while often produced under competitive conditions, have no individual costs which can be determined. However, this situation does not result in any very serious modification of our long-run, competitive value theory. If two articles are produced under conditions of joint costs, their combined prices in the long run under competitive conditions must equal the single and indivisible cost of producing them or production cannot be maintained. Neither article individually tends to have a price equal to cost in the long run since neither one has a determinable cost. Therefore the prices of the individual products tend to depend in the long run on the relative strength of the demands for these individual products. That is, if many people in this country should become vegetarians and abstain from meat consumption, while the demand for shoes and other leather products went on as at present, the price of hides would be expected to make the major contribution, and that of carcasses the minor contribution, toward covering the cost of producing steers in the long run. Similarly a great increase in the demand for butter substitutes, cooking oils, and other things which use cottonseed oil as an ingredient, while rayon goods were gradually replacing cotton textiles, would in the long run increase the relative contribution made by the price of cottonseed toward the cost of raising these joint products, and diminish that made by the price of cotton.

The Effects of Changes in Demand. A change in the demand for one of two joint products, while that for the other joint product remains the same as before, introduces some complications into the discussion of price determination under conditions of joint costs. The general theory covering this situation states that a change in the demand for one joint product, other things being equal, tends to produce a like change in its price but an inverse change in the price of the other joint product. That is, if the demand for cotton increases, the price of cotton tends to increase also in the long run, while that of cottonseed tends to decline. If the demand for cotton decreases, the price of cotton tends to fall in the long run while that of cottonseed tends to rise. However, it is another thing to explain the process by means of which these changes take place.

If the demand for cotton increases, the price of cotton tends to rise abruptly in the short run. The combined prices of the cotton and cottonseed now exceed the cost of producing the joint products. If this profitable situation continues for any great length of time, an expansion of the productive capacity and normal annual output of the industry tends to take place under competitive conditions. Since the industry is one of increasing cost, this expansion tends to increase the long-run minimum average cost

of production per unit of the joint products, but the combined prices of the products will tend once more to equal cost. The expansion of cotton production cannot occur without also increasing the production of cottonseed. The demand for cottonseed has not increased, while more cottonseed than formerly is produced and must be sold. The price which the cottonseed commands will therefore be lower than that which prevailed under the old long-run adjustment. Since the combined prices of cotton and cottonseed must now be higher than formerly in order to cover cost, while the price of cottonseed is lower than its former price, the price of cotton, though it will be lower than the short-run price under the increased demand, must be higher than it was under the old long-run adjustment.

The effects of a decrease in the demand for either joint product can be explained by a similar process of reasoning. It is also possible by means of diagrams to illustrate all these conclusions with regard to price determination under conditions of joint costs, but these diagrams are so complicated that their probable usefulness to the beginning student of economic principles is relatively small. Since the purpose of presenting diagrams is to assist in the understanding of the principles of price determination, it seems likely that the diagrams may be safely omitted in cases where they are more complicated than the theories which they illustrate.

Joint Demand. There is said to be a joint demand for the various things which go into the production of a finished article, and also for various finished products which are usually consumed together. The demand for houses really constitutes a joint demand for bricks, lumber, plumbing and electrical fixtures, the labor of workers in the various building trades, and other things. There is also a joint demand for golf clubs and golf balls on the part of persons who follow that ancient and royal pastime, and one for automobiles and gasoline. The prices of goods which are jointly demanded ordinarily tend to move in opposite directions when some change in supply upsets the existing relationship. If the production of golf balls becomes greatly restricted and the price rises to an average of about $2.00 per ball, the playing of golf will be checked (under given demand and income conditions), the demand for golf clubs will decrease, and the price of golf clubs will fall. Gasoline is used for other purposes than the operation of motor vehicles, but much of it is used in automobiles. If the price of automobiles increases abruptly and the quantity produced and sold declines, the operation of automobiles will eventually fall off, the demand for gasoline will decline to a considerable extent, and the price of gasoline will tend to decrease.

A similar situation prevails in connection with the prices of the materials which go into the making of a finished product. Under given methods of production, if the production of lumber is restricted and its price rises considerably, the amount of housebuilding will tend to be re-

duced, the demand for other things used in housebuilding will tend to decline, and the prices of these things will tend to fall. This discussion, of course, assumes that, as the production of lumber is restricted, the demand schedule for new houses and the supply conditions for plumbing fixtures, bricks, plaster, labor, and other things used in housebuilding remain unchanged. Under the same demand schedule, the prices of new houses tend to be higher than before, while the prices of the things other than lumber which are used in housebuilding tend to be lower than before. The houses sell for prices well above the sum of the costs of these other things used in housebuilding, and the difference between these amounts limits the extent of the possible rise in the price of lumber, if lumber is indispensable in housebuilding. Therefore the price which can be commanded by anything used in the production of a finished article is limited, for any given output of the finished article, by the margin between the price per unit at which such an output of the finished article will be purchased by buyers and the sum of the costs, per unit of the finished article, which must be paid for the other things necessary to produce it.

Obviously the variable elements in this situation are very numerous. If the production of lumber is restricted, as assumed above, its price will tend to rise much more if lumber is absolutely necessary to housebuilding than if it is not. If a small decrease in the demand for the other things used in housebuilding will result in a considerable drop in their prices, the restriction of lumber production will tend to raise the price of lumber much more than if the reverse is true. If the demand for new houses is quite inelastic, the restriction of lumber production will tend to raise its price more than it will if the demand for new houses is elastic. Finally, if the cost of lumber plays only a relatively small part in the total cost of housebuilding, the restriction of its production would raise its price more than it would if the reverse were true.

The discussion of the relationships between the prices of goods which are jointly demanded has assumed that some change on the supply side has caused the price of one such product to increase and in such cases the prices of the other products tend to decrease. However, if one of two jointly demanded products experiences a rise in price because of an increased demand, the price of the other product will tend to rise also. Increase the demand for golf clubs and more clubs may be purchased than formerly though the price rises. The increased use of golf clubs leads to an increased demand for golf balls, assuming no changes in skill on the part of the users of golf clubs, and the price of golf balls rises also.

Composite Supply. The situation known as composite supply exists when a given demand can be satisfied quite well by any one of a number of products. These various commodities or services compete with each other to satisfy the demand and are often called competing goods. For

traveling between New York and Chicago a composite supply of transportation is found, since one's desire for movement between these points can be satisfied by the use of railroads, bus lines, air lines, or private automobiles. Similarly, the desire for a bread spread can be satisfied by using butter or oleomargarine. The prices of competing goods which make up a composite supply tend to move upward and downward together when something affects the supply of one good. If a shortage of butter causes its price to rise, some consumers turn to oleomargarine, the demand for the latter product increases, and its price also tends to rise. When butter becomes plentiful and its price falls, some consumers turn from oleomargarine to butter, the demand for oleomargarine falls, and its price declines. In similar fashion, a decrease in railroad fares between two points will usually lead to a decline in bus fares as well. On the other hand, if the price of butter rises because of an increase in demand rather than a decrease in supply, the use of butter may be great even at a higher price than formerly, fewer people will use butter substitutes, and the price of oleomargarine tends to fall.

Composite Demand. When a certain finished good or factor of production is demanded for use for a number of different purposes, the situation is known as one of composite demand. For example, land of a given grade for agricultural purposes may be in demand to raise corn, soybeans, and wheat, while land of a given grade for site purposes in the downtown section of a city may be demanded as sites for theaters, office buildings, retail stores, and hotels. If the demand for the agricultural land to raise soybeans increases and the rent of land in that use increases, there will be less of the land than formerly for raising corn and, if the demand for land for this purpose remains the same, the rent of cornland will also tend to rise. Thus, as we saw in connection with the subject of opportunity costs, land of a given grade will tend to command the same rent in each of its alternative uses in the long run, and a rise in its rent in any one use will tend to be associated with a rise in its rent in other uses. Moreover, if the price of soybeans rises considerably, the price of corn will tend to rise too, other things being equal, because the rise in the price of soybeans will tend to set in motion the process described above. If the production of two economic goods requires the same agents of production, a rise in the price of one, which causes agents of production to be transferred to its production, tends to decrease the agents available to produce the other good, and, under given demand conditions, the price of the other good tends to rise also.

QUESTIONS AND PROBLEMS

1. Why does the economist maintain that the price of an economic good tends to equal minimum average cost of production for all the firms in a competitive industry in the long run? Explain.

2. "The Law of Long-Run Competitive Price is clearly worthless because business enterprisers will not remain in business unless they can sell their goods for prices which yield a margin of profit over and above cost of production." Show whether you agree.

3. "When an enterpriser has a stock of goods on hand, he will ordinarily have to sell them within a comparatively short period of time for whatever price can be obtained, whether or not this price is sufficient to cover cost of production. This disproves the economist's contention that the price of an economic good tends to equal average cost of production in the long run under competition." Do you agree? Explain.

4. "The long-run price of an economic good is determined by its cost of production under competitive conditions." Show whether you agree.

5. "The price of wheat last year was $1.50 per bushel. This shows that the cost of production of wheat last year was also $1.50 per bushel." Do you agree? Explain.

6. "All industries are industries of increasing cost in periods of rising prices." Show whether you agree.

7. "Almost any industry will be one of decreasing cost over the first few years of its life." Do you agree? Explain.

8. "The use of the mechanical cotton picker in cotton raising is likely to change the industry from one of increasing cost to one of decreasing cost." Show whether you agree.

9. In explaining the determination of long-run prices, economists suggest that industries may be divided into three classes: i.e., industries of increasing, decreasing, and constant cost. Explain exactly what we mean in speaking of these three types of industries.

10. There are at least three senses in which an individual industry might be described as "an industry of decreasing cost." Explain these meanings of the term and show which one is used by the economist in connection with long-run price determination.

11. "It is a mistake to decide that an industry is one of decreasing cost merely because its firms experience lower cost as they increase their rate of operation with fixed productive facilities from part capacity to normal output." Explain.

12. Explain price determination in the long run under competition under conditions of increasing cost (or constant cost, or decreasing cost).

13. "The effect of technological changes is on the *level* of entire long-run cost and supply curves and not on their *slope*." Explain.

14. Distinguish between joint products and by-products.

15. "The theory of competitive price determination in the long run has to be modified only slightly to take care of price determination for joint products." Explain.

16. If the United States should resume the heavy importation of raw silk from the Far East in the postwar period, how would you expect the price of cottonseed meal to be affected? Explain.

17. Since steel and rubber are jointly demanded for the production of auto-

mobiles, how would you expect a considerable decline in the price of steel to affect the prices of rubber and automobiles? Explain.

18. How would you expect a greatly increased consumption of pork in the United States to affect the price of shoes?

19. Distinguish between joint supply and composite supply.

20. "When articles are jointly demanded, if the price of the one good rises, that of the other will always fall." Show whether you agree.

See References for Further Reading at the end of Chapter XVI.

6. "In disposing of a perishable fixed stock the monopolist has a distinct advantage over any one seller in a competitive market." Show whether you agree.

7. "By disposing of his product for greatest total receipts, the monopolist assures himself of the best possible financial result in the short run." Explain.

8. "In the intermediate period, the monopolist secures the best possible financial result by expanding output and sales to the point at which marginal revenue equals marginal cost." Do you agree? Explain.

9. "By producing the output for which marginal revenue is equal to marginal cost, the monopolist may make a profit, sustain a loss, or exactly cover average cost of production, in the intermediate period." Explain.

10. Discuss the monopolist's prospect of increasing his profits through advertising.

11. Explain price determination in the long run under monopoly.

12. "In the long run, the monopolist would adjust the size, productive capacity, and output of his industry so that the marginal revenue derived from the sale of the product would equal the long-run marginal cost of producing it." Show whether you agree.

13. "In the long run the effect of the expansion of a monopolized industry on the prices of the productive factors is less important than its effect on the internal efficiency of the monopoly enterprise." Explain.

14. "Monopoly prices are always higher than competitive prices." Show whether you agree.

15. "A monopolist is very sure to make profits in the long run." Do you agree? Explain.

16. Summarize the determination of prices under monopoly.

17. "A monopolist will make larger profits by selling his good on the basis of class price than by charging any single price." Explain.

18. "The existence of class price depends more upon the imperfection of demand than upon conditions of monopoly on the side of supply." Show whether you agree.

19. "The price charged for a good under duopoly may fall anywhere in the range from the price which would prevail under complete monopoly to that which would prevail under competition." Explain.

20. "There are many fields of production in which there are only a few sellers, but only a few of these cases represent pure conditions of oligopoly." Explain.

See References for Further Reading at the end of Chapter XVI.

XVI

The Determination of Prices under Noncompetitive Conditions (*Continued*)

The final phase of our theory of price determination deals with the market situation known as monopolistic competition. The distinguishing characteristic of monopolistic competition is found in the fact that, by means of physical differences, patented features, and other devices mentioned in Chapter X, the products of the various firms in an industry are differentiated so that the firms do not sell exactly the same economic good. The number of firms in the industry under monopolistic competition is supposed to be just as great as under competition. The firms operate independently and do not enter into agreements or conspiracies to control the market, although the actions of the individual firm are more important for the other firms than would be the case under competition. Finally, the firms are supposed to have adequate information concerning demand and market conditions in general.

SHORT-RUN PRICES UNDER MONOPOLISTIC COMPETITION

The Demand Situation. The demand for the product of the individual firm under monopolistic competition is different from that which prevails under either competition or monopoly. Under monopoly the demand for the product of the individual firm is the entire demand for the economic good in question in the market, but under monopolistic competition there are other producers of the same general kind of product, and the demand for the product of any one firm is only a partial demand. Under competition the demand for the product of the individual firm is represented by a horizontal line parallel to the *OX* axis in a diagram, and this indicates that the individual firm can sell any amount of the product which it chooses at the prevailing price, but that it cannot sell at a higher or lower price. This type of demand results from the fact that under competition there are many producers of exactly the same economic good, but this condition does not prevail under monopolistic competition. Since the products of the individual producers are differentiated, the demand curve for the product

of the individual firm will move downward from left to right in a diagram, indicating that the quantity of this firm's product which buyers would purchase would vary inversely with the price charged for it, in a given market at a given time.

The exact nature and position of the demand curve for the individual monopolistic competitor's product depend upon the variety of the general good which this firm decides to produce, the varieties which other firms decide to produce, the prices which other firms charge, and other factors. However, the demand for the one firm's product is likely to be predominantly elastic, and this tends to be true, as we noted in Chapter XI, even where the demand for the product in general is predominantly inelastic.

Price Determination under Perishable Fixed Stock. The individual monopolistic competitor, like the individual seller under other conditions of the market, will have a fixed stock of his product on hand in the short run, and the commodity may be highly perishable. In this condition of perishable fixed stock it is very difficult for the individual monopolistic competitor to decide what to do about the price of his product unless he knows something about the situation of the other monopolistic competitors. That is, as noted above, there is no definite demand for one monopolistic competitor's product unless his variety of the product, other firms' varieties of the product, and the prices charged by other firms are taken as given. Of course, the nature of the products of all monopolistic competitors is definitely determined in the short run, but there may still be a question as to whether the other firms are also caught with a stock of highly perishable goods at the moment and, if so, what they will do about it.

Given the necessary assumptions concerning the situation of other firms and the prices they will charge for their products, the demand for the one monopolistic competitor's product would take definite shape. Under it he would try to maximize the total receipts from the sale of his goods, whether this involved selling all of the fixed stock or only part of it. Since the demand for his product is predominantly elastic, he would maximize his total receipts by selling all of the perishable fixed stock unless the stock were of very large size. As in the case of the monopolist, selling his perishable fixed stock so as to secure the greatest possible total receipts would assure the monopolistic competitor of the best possible financial result, whatever the total costs of the goods may have been and whether the result were a profit or a loss. In actual practice, the best decision of the monopolistic competitor as to the price at which to sell his perishable fixed stock may prove to be incorrect. That is, in adjusting his price so as to obtain maximum total receipts he may cause the other monopolistic competitors to change their prices, whether or not they are operating under conditions

of perishable fixed stock, and the result would be a change in the market demands for the one monopolistic competitor's product at various prices.

Price Determination under Ordinary Short-Run Conditions. Under ordinary short-run conditions, the individual seller under monopolistic competition will find on his hands a fixed stock of an economic good which is not highly perishable and he will, like the full monopolist, try to dispose of this stock over the several available market days in such a way as to maximize the total receipts from its sale. However, he might experience considerable difficulty in attaining this objective, for any adjustment of his price for the purpose of maximizing total receipts would be most likely to stir up price changes on the part of the other monopolistic competitors. If we assumed that the other monopolistic competitors would not change their prices, the principles governing the activities of the monopolist in the ordinary short-run situation would apply fully to the monopolistic competitor. That is, he would sell on a given day some quantity between the limits of no units and the number of units which would give him maximum total receipts on that day. The exact quantity to be sold on a given day would depend upon the individual monopolistic competitor's estimate with regard to market conditions of the present and near future, and the price would be that at which buyers would purchase the quantity of product in question. The monopolistic competitor, like the monopolist, may or may not make a profit under ordinary short-run conditions, but he cannot do better than to sell his fixed stock for greatest possible total receipts.

INTERMEDIATE-PERIOD PRICES UNDER MONOPOLISTIC COMPETITION

The Cost Situation. In the intermediate period the individual firm under monopolistic competition, like the individual firm under any other conditions of the market, has a fixed amount of plant and equipment (as well as fixed amounts of certain other productive factors) and a fixed productive capacity, so that changes in output can result only from speeding up or slowing down the rate at which the fixed productive facilities are operated. In this period the monopolistic competitor has some costs which remain fixed in total amount regardless of changes in output, and other costs whose total amount varies directly with output. The distinction between average and marginal costs is also important for the monopolistic competitor in this period, and we may expect these types of cost to behave in their usual fashion as the monopolistic competitor operates his fixed productive facilities at varying rates, although the exact position and slope of his marginal and average cost curves in a diagram would depend upon the particular variety of the general product which he decides to produce. Assuming that the individual monopolistic competitor has chosen to produce a particular variety of the general product, we take the cost curves in Figure 31 to represent the behavior of his marginal and average costs.

Price Determination for the Individual Firm. In the intermediate period as in the short run, the demand for the individual monopolistic competitor's product takes definite shape only if we assume that he has settled upon a particular variety of the product to produce and sell, and that other

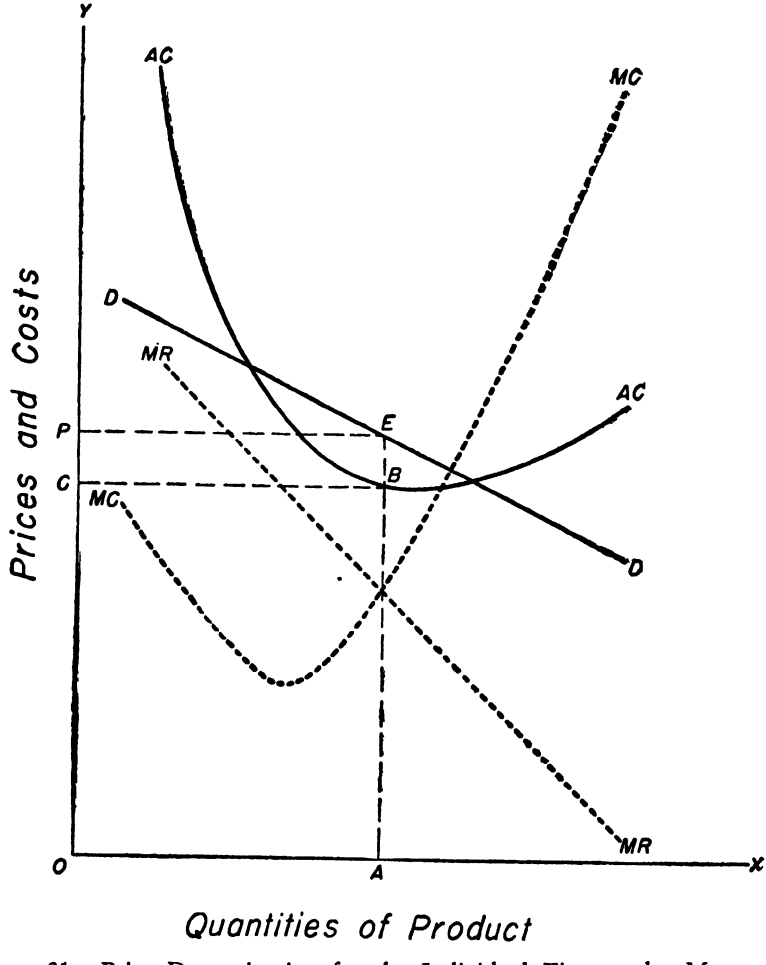

FIGURE 31.—Price Determination for the Individual Firm under Monopolistic Competition in the Intermediate Period

monopolistic competitors have decided upon both the particular varieties of the product which they will produce and the prices which they will charge. Having made these necessary assumptions, let us suppose that the demand for the individual monopolistic competitor's product is as represented by the line *D* in Figure 31. Under this favorable condition of

demand, there are clearly several outputs which the monopolistic competitor could produce and sell at a profit; that is, there are several outputs which would sell at a price in excess of average cost per unit of product. How can the monopolistic competitor determine the output which he should produce and sell, and the price which he should charge?

The monopolistic competitor, like the individual enterpriser under any other condition of the market, is anxious to obtain the best possible financial result from the operation of his business. With this in mind, he will be willing to increase his output in the intermediate period as long as each additional amount of product increases his total receipts more than it increases his total costs; that is, as long as marginal revenue is in excess of marginal cost. The actual output selected will be (as nearly as possible) that at which marginal revenue and marginal cost are equal. To stop short of this output would be to fail to produce some units of product which would add more to revenue than to cost, while to go beyond this output would involve producing some units which would add more to cost than to revenue. Thus, under the demand and cost conditions shown in Figure 31, the monopolistic competitor would produce an output of *OA* units per year and sell it for a price of *OP* per unit. Total receipts would be *OAEP* and total costs, based on an average cost of *AB* or *OC* per unit, would be *OABC*. The monopolistic competitor's total profit would be *CBEP*, and this profit would be the largest one possible under the given conditions of demand and cost.

As was true in the case of monopoly, under other and less favorable conditions of demand the demand curve for the individual monopolistic competitor's product might touch his average cost curve at only one point, or might even miss the average cost curve altogether. In the first case, the monopolistic competitor would produce and sell his good for a price equal to average cost of production per unit, while in the second case he would find it necessary to operate his business at a loss. In both cases, however, the most favorable output in the intermediate period would be that at which marginal revenue was equal to marginal cost. Diagrams representing these situations are not shown here, but they would be essentially similar to those presented in connection with the determination of monopoly price in the intermediate period.

By changing some of our assumptions we could secure other variations of the problem faced by the individual firm under monopolistic competition in the intermediate period. If we assumed that all other firms had determined both the varieties of the product which they were going to produce and the prices which they were going to charge, while our individual firm had determined its selling price but not the variety of the product which it was going to produce, other diagrams could be constructed to show how

the firm would determine the most favorable variety of the good to produce. In similar fashion, other diagrams could be produced by assuming that the prices charged by all other firms, instead of being already determined, tended to vary with the price charged by our individual firm, or by considering the possible effects of advertising outlays on the demand for the product of the individual firm. However, such complications in diagraming seem hardly necessary for the beginning student of economic principles.

The Possible Effect of Advertising. It would be somewhat more difficult for the monopolistic competitor than for the monopolist to decide whether it would be worth while to make extensive outlays for advertising. The monopolistic competitor would have to estimate that, by incurring advertising costs in varying amounts, he could bring about varying increases in total sales if the other monopolistic competitors continued to operate their businesses as before. By deducting each possible total cost of advertising from the total receipts which would be derived from sales in each case, he could determine the amount of income which would be left over for his production department in each case. Then he would have to see how his ordinary costs of production would be affected by expanding output to various extents under different amounts of advertising. Finally, he would have to reckon with the probability that the other monopolistic competitors would not merely stand idle while he was puffing his own wares. That is, he would have to estimate what the other monopolistic competitors would do about advertising if he undertook an advertising campaign, and try to figure out how his own advertising campaign would be affected by those undertaken by other sellers. With all these things in mind the individual monopolistic competitor would have to decide to what extent, if at all, it would be worth while for him to undertake an advertising campaign for his product.

Difficulties in Dealing with a Whole Industry under Monopolistic Competition. It is rather difficult to deal with price determination for all the firms in an industry under monopolistic competition in the intermediate period. There may be significant variations in prices and products from one firm to another and the variations in prices and products may not be at all evenly spaced from firm to firm. The demand curves for the products of different firms may also show a considerable variation, and the slopes and positions of the cost curves may differ greatly from one variety of this product to another. Moreover, the customers of one firm may be much more securely attached to it than the customers of other firms are to them. In view of these difficulties we shall not discuss the question of price determination for a whole industry in the intermediate period under monopolistic competition.

LONG-RUN PRICES UNDER MONOPOLISTIC COMPETITION

Industrial Expansion and Contraction. Individual firms which make profits in operating their businesses in the intermediate period under monopolistic competition could continue to make profits over a considerable period of time if most other firms in the same industry were breaking even or making losses. However, there would probably be a tendency for the less successful firms to imitate the products of the profit-making firms more closely as time went on so that they would come to have approximately the same costs. If profits or losses were made by all or practically all of the firms in an industry under monopolistic competition, the whole industry would probably expand or contract in size and productive capacity in the long run. Under monopolistic competition, as under other conditions of the market, the long-run period is one in which changes can be made in industrial size and capacity and not merely in the rate at which fixed productive facilities are operated.

As under competitive conditions, it is likely that the expansion or contraction of a whole industry under monopolistic competition would occur by means of the entrance or exit of firms and not by means of the expansion or contraction of the size of the firms already in the industry. Under any given demand situation, the individual firms are likely to develop to what they regard as optimum size, with regard to the productive agents at their disposal, in the long run. If this is true, the firms are not likely to expand to larger size even though the market situation for the industry as a whole becomes very favorable and large profits are being made, for such expansion would place the firms at a disadvantage in comparison with other firms of optimum size which might come into the industry to produce approximately the same product. In similar fashion, under an unfavorable market situation for the industry as a whole, it is more probable that some firms will leave the industry than that all firms will contract to some inefficient size short of the optimum.

Increasing, Decreasing, and Constant Cost. The expansion of an industry by means of the entrance of firms under monopolistic competition would differ from the expansion of a competitive industry in that the new firms which entered the industry would not produce exactly the same product as the firms already in the industry. However, in spite of product differentiation, it is probable that the new firms would require almost exactly the same agents of production as those used by the old firms. Hence the expansion of the industry by means of the entrance of firms might result in an increase or a decrease in the costs of certain productive agents to all the firms, or by chance might leave the costs of all productive agents unchanged. In converse fashion, the contraction of the industry by means of the departure of firms might result in a decrease or an increase in the costs

of certain productive agents to the surviving firms or might leave the costs of all productive agents unchanged. The conclusion is, then, that an industry under monopolistic competition might operate under conditions of increasing, decreasing, or constant cost in the long run.

Price Determination. In reaching a conclusion as to the probable level of prices in an industry operating under monopolistic competition in the long run, we start with an intermediate-period situation in which the demand is highly favorable for all or most of the individual firms in the industry. In other words, all or most of the firms are in a position similar to that of the firm whose affairs were illustrated in Figure 31. Almost every firm has a choice of profitable outputs, and it attempts to maximize profits by selecting that output at which marginal revenue is equal to marginal cost. In spite of the profits which are made, there is no reason for any firm to produce a larger output per year, since the marginal cost of any additional units of product would be greater than the marginal revenue which they would furnish. Similarly, if any firm produced a smaller output per year, it would be failing to produce some units which would furnish a marginal revenue in excess of their marginal cost. The situation is therefore one of stable equilibrium *for the intermediate period,* which is too short to allow new firms to come into the industry.

In the long run the large profits made by the individual firms already in the industry would be expected to attract new firms to the industry. If the industry were one of constant cost, the entrance of new firms would have no effect on the average cost curve of each firm in the industry. If the industry were one of increasing or decreasing cost, the expansion of the industry and entrance of firms would raise or lower the average cost curve of each individual firm. Whatever might happen to average cost, the demand for each firm's product would be certain to change as the industry expanded. That is, each firm would experience a decrease in demand as the unchanged total purchases of the goods were divided among a larger number of firms.

What would be the final result of the expansion of the industry, the changes in demand for the products of the individual firms, and the possible changes in the average costs of the firms? An industry composed of numerous firms cannot be said to be in long-run equilibrium so long as the individual firms make net profits. Therefore the expansion of the industry would stop and the industry would be in long-run equilibrium only when there was no output that each individual firm could produce at a profit and only one output that each firm could produce and sell at a price equal to average cost. (This would be represented by a diagram in which the demand curve for each firm's product would be tangent to its average cost curve at some point or other.) Our net conclusion is that profits would be eliminated in the long run in an industry operating under monopolistic

competition, and that each firm would get a price for its product which only equaled average cost of production, although this cost of production might not be the same as that which would be equal to price under full competitive conditions.

The same conclusion would be necessary in the end if we started our discussion with an intermediate-period situation in which demand was highly unfavorable for the individual firms in the industry under monopolistic competition. Suppose, for example, that the demand for each firm's product was so weak that there was no output which the firm could produce and sell at a price per unit even equal to average cost of production. It would presumably still be worth while for each firm to continue to operate since it would minimize its losses by so doing, and the output at which losses would be least would be that at which marginal revenue was equal to marginal cost. There would be no reason for any firm to produce a smaller output since the result would be a failure to produce some units which would furnish a marginal revenue in excess of their marginal cost. Conversely, there would be no reason for any firm to produce a greater output, since the marginal cost of any additional units produced would be greater than the marginal revenue which they would furnish. The existing situation would therefore be one of stable equilibrium *in the intermediate period,* which is too short to permit firms and agents of production to leave the industry.

In the long run the considerable losses made by the individual firms would be expected to cause some firms to leave the industry. If the industry were one of constant cost, this departure of firms would have no effect on the average cost curves of the surviving firms. If the industry were one of increasing or decreasing cost, the departure of firms would lower or raise the average cost curve of each surviving firm. Whatever might happen to average cost, each firm would experience an increase in the demand for its product as given total purchases were divided among a smaller number of firms.

The contraction of the industry, changes in demand for the products of the individual firms, and possible changes in the average cost of production of the firms would all stop when the industry reached a position of stable equilibrium in the long run. This position would be reached when each firm found that there was no output which it could produce and sell at a profit but there was one output which it could produce and sell at a price equal to average cost. As before, this situation would be represented by a diagram in which the demand curve for each firm's product would be tangent to its average cost curve at some point or other. Thus losses would be eliminated in the long run in an industry operating under monopolistic competition and price would be equal to average cost for each firm.

Price under Monopolistic Competition Crossed with Oligopoly. Our conclusion that price tends to equal average cost of production for each firm under monopolistic competition in the long run depends upon our previous assumption that the firms in such an industry are as numerous and small as they would be under competition. In this situation the industry as a whole will be expected to expand or contract, by means of the entrance or exit of firms, in the face of continued profits or losses. If, however, as in many practical situations, monopolistic competition is crossed with oligopoly and there are only a few large sellers of differentiated products, a quite different conclusion is necessary. That is, if an industry is controlled by a few large firms, so that a tremendous aggregation of capital funds would be required to set up new firms capable of competing with those already in the industry, entrance into the industry will be relatively difficult and considerable profits may be made year after year by the firms already in the industry without causing any new firms to enter, even in the long run.

Again, if the firms in the industry are few and large, even a considerable period of adversity may not cause any firms actually to leave the industry. The firms may "fail" in the sense of being unable to make ends meet financially but, instead of leaving the industry altogether, they will be likely to reorganize in such a way as to lower some of their fixed costs. Bondholders who own large quantities of bonds bearing high rates of interest may be asked to accept smaller quantities of bonds with lower rates of interest, salaries of permanent officials may be cut, and so on. In this way the average costs of the firms may be brought down so that, even under existing unfavorable conditions of demand, it will be possible for them to produce and sell some output at a price which covers average cost.

Summary. In summary we may say that the individual monopolistic competitor acts as nearly like a monopolist as possible. In the short run, when he has a fixed stock of his product on hand, he tries, within the limitations of his position, to maximize total receipts from the sale of his good. If successful, he is assured of the optimum financial result in this period—a maximum profit or a minimum loss. In the intermediate period, when the monopolistic competitor has an establishment of fixed size and productive capacity which may be operated at various rates, he will try to gauge his output so that marginal revenue from sales will be equal to marginal cost of production. This output and the price at which it can be sold may cause the monopolistic competitor a loss, bring him a profit, or by chance merely allow him to recover total costs of production, but in any case the financial result will be one which cannot be improved upon within the limits of the intermediate period. In the long run, his firm having been developed to optimum size, the problem of the *individual* monopolistic competitor remains essentially the same as in the intermediate period. In all periods of time it must be emphasized that the demand for the product

of the monopolistic competitor is not a simple function of the various prices which he might charge but depends also on the variety of the general product which he produces and sells, the varieties produced and sold by other monopolistic competitors in the same general field, and the prices charged by the other firms for their closely similar products. This fact greatly complicates the problem of the individual monopolistic competitor in deciding matters of output, sales, and price.

In the long run, changes in the size and productive capacity of a whole industry under monopolistic competition are likely to occur by means of the entrance or exit of firms rather than by means of the expansion or contraction of the size of individual firms. Under normal conditions of monopolistic competition the whole industry will expand or contract in size and productive capacity, in the face of widespread profits or losses within the industry, until price tends to settle at the level of average cost of production for the individual firms and profits or losses are eliminated. Such a price, when attained, will be higher, and the output of the whole industry will be smaller, than would be expected under competitive conditions. If the firms in the industry are few and large, so that conditions of oligopoly are crossed with those of monopolistic competition, the conclusions as to the nature of the long-run adjustments which will occur in the industry need to be modified considerably.

THE VALIDITY OF THE THEORY OF VALUE IN NONCAPITALISTIC SYSTEMS

Demand Theory under Socialism. Our discussion thus far has attempted to describe the processes of price determination as they exist and operate under capitalism, or as they would exist and operate under idealized conditions under capitalism. We shall now consider briefly the extent to which this explanation of price determination would be applicable, if at all, under other types of economic systems. Under socialism, since individuals are expected to have money incomes and are supposed to be free to spend these incomes as they desire for the various economic goods which are made available, much of our analysis of the demand for economic goods might be salvaged. The Law of Diminishing Utility, for example, might work as well under socialism as under capitalism. For the individual consumer under either system the desirability of acquiring a second suit of clothes of a given kind would be less than that of the first suit, and the desirability of a third suit would be less than that of the second.

Since the Law of Diminishing Utility would be working, and since differences in desires and, to some extent, in incomes would prevail among individuals under socialism, the Law of Demand would probably be valid. That is, buyers would be expected to purchase larger quantities of an economic good at low prices than at high prices at a given time and place. Under socialism as under capitalism the demands for some goods might be

predominantly elastic, and those for other goods predominantly inelastic. However, an increase or a decrease in the demand for an economic good would not be likely to affect production in the same way under socialism as under capitalism.

Prices and Costs under Socialism. With the exception of the theory of demand, our analysis of price determination would be useless for the most part in connection with a socialistic economic system. Take, for example, the principle which holds that the price of an economic good, in the long run under competition, tends to equal the minimum average cost of production of all the enterprises in the industry. There is no reason to expect the price per unit of an economic good to equal cost of production under socialism.

In the first place, it would be very difficult to determine the cost of production of any good under socialism. For all practical purposes society as a whole would be the only business enterpriser and would both own and use the land and capital of the system. With society as a whole, operating through the planning commission, functioning as the sole supplier and demander of land and capital, there would be no "market" for these agents. The planning commission might well charge itself with certain amounts of rent and interest for the land and capital used in various branches of production, but such charges, being quite arbitrary and useful primarily for accounting purposes, might differ widely from what the costs of these agents would be in competitive markets.

Wages of various kinds would constitute about the only money expense of production in the ordinary sense under socialism. That is, wages would be a real cost of production, and not merely an accounting cost, since their payment to individuals would be necessary to the continuation of production. However, wage differentials would be juggled by the planning commission to achieve an appropriate distribution of labor among industries and occupations as well as to reward the workers for their contributions to production, and they might differ considerably from those which would result from the competition of workers and private employers in a free market. Thus under socialism the cost of any economic good would be whatever the planning commission said it was. Even the prices of finished commodities and services would be controlled, directly or indirectly, by the planning authority, and would not be determined by competition between buyers on the one hand and private sellers, organized under competitive conditions or those of monopoly or monopolistic competition, on the other.

In the second place, after the prices and costs of economic goods had been determined in arbitrary fashion, there would be no necessary relationship between them. A given good might be selling at a price well above its alleged cost of production per unit and the planning commission might decide to decrease its production or even discontinue it altogether, on the ground that the good catered to wants which might better be left unsatis-

fied or that the production of something else was more important to society as a whole. Another economic good might be selling at a price far below its alleged cost of production per unit, and the planning authority might decide to double or treble its production as soon as possible, because it was thought that society would gain from widespread consumption of the good. Of course, in either case the planning commission could make arbitrary price equal arbitrary cost if it so desired, but there would be no reason why it would have to do so.

As soon as prices and costs are arbitrarily determined by the same agency, the relationship between them loses significance from the point of view of controlling production. That is, the planning commission, having set the price of shoes of a certain grade at $5 and their cost at $4, would be very silly to decide to expand the production of the shoes on the ground that a "profit" of $1 per pair was being made. Again, as soon as production does not need to be affected by the relationship between price and cost, any compelling tendency for the price of an economic good to equal cost of production per unit disappears. Under socialism alternative or opportunity costs are the only costs which the planning commission would have to consider very seriously. That is, with fixed total amounts of the productive agents available, the cost of getting more of any particular kind of economic good is found in the necessity of getting along with smaller quantities of other economic goods than could have been enjoyed if the production of the first good had not been increased. If agents of production are used in one field of production, they cannot be used in other fields of production at the same time.

Profits and Losses under Socialism. But, it may be objected, will not a money loss which is taken on one economic good have to be offset by a money profit made on some other economic good or goods so that, for the system as a whole, prices tend to cover costs? Surely the entire economic system could not be operated at a loss under socialism? The answer is that it may be convenient if costs and prices (or income) balance for the system as a whole, but a failure to operate on a "profitable" basis, or even on a cost basis, will not upset or interfere with the operation of the economic system. The concepts of profit and loss in terms of money are important only to individuals and only in systems in which industries and businesses are privately owned and operated. If the individual in our capitalistic system makes a profit, he increases his relative command over the real income of the system, while the opposite result occurs when he sustains a loss in the operation of his business. But society as a whole under socialism cannot increase its total real income by one iota by making a money profit on the operation of the economic system, nor can it reduce this real income at all by sustaining a money loss. Thus money profits and losses are of no real significance in a socialistic system.

Wage payments constitute about the only necessary money expense under socialism, and the operation of the whole system at a loss would apparently mean that the total sum paid out as wages is greater, in a given period, than the amount of money necessary to take all available economic goods off the market at the prices which the planning authority sets on them. The only effect of this situation would be that the individuals of the system as workers and consumers would have money left over in their pockets after buying all the finished goods and services on the market. Regardless of the amounts of their money income, they cannot buy more goods than are produced for them nor can they bid up the prices of the available goods, since these prices are fixed by the planning authority. While the individuals of the system might be dissatisfied at having certain amounts of money income which could not be spent for anything, the operation of the system at a loss would not be likely to have any very serious results.

It is not easy to see how the whole economic system could be operated at a money profit under socialism, for this would involve selling the available commodities and services for a total sum of money greater than the amount paid out as income to the consumers of the system. In any case, since the planning authority would control wage payments, the quantities of finished goods on the market, and the prices to be charged for these goods, the authority should find it possible to achieve for the system as a whole a balance between the total payments of money income to the workers and the total payments to be received from the sale of finished goods at established prices. Thus it should be possible, if desired, to "come out even" on the operation of the whole system, even though individual goods sell for more or less than their nominal costs of production.

Monopoly Price under Socialism. Since most fields of production would be operated as governmental or social monopolies under socialism, it might be thought that our principles of monopoly price determination would be of some significance in that system. However, although individual governmental monopolies under capitalism sometimes operate so as to obtain the greatest possible total profit, it is unthinkable that all lines of production would be operated in that way under socialism. Production under socialism is supposed to be controlled on the basis of social need and in such a way as to satisfy the wants of the consumers as a whole as completely as possible, while the operation of all lines of production on monopoly principles would result in a great restriction of production and real income. Society as a whole operates an economic system "profitably" only when it succeeds in accomplishing whatever productive results it sets out to achieve.

Production and price determination under conditions resembling monopolistic competition, with, from the social point of view, its largely unnecessary product differentiation, would apparently be entirely incom-

patible with the aims and principles of socialism. Our final conclusion is, therefore, that our description of price determination, with the exception of demand principles and the principle of opportunity costs in a physical sense, would not be valid under socialism. Under socialism society as a whole produces and consumes what it wants to produce and consume, as represented by the decisions of the planners, and prices and costs, as expressed in terms of money, are used largely for convenience in planning, carrying out the plans, and measuring the results. Prices and costs are the servants of such an economic system and not its masters.

Value Theory in Relation to Communism. Our description of price determination would be even less useful for a communistic system. Since communism hopes to avoid the use of money and prices altogether, no theory of price determination could be significant in that system. And since individuals would presumably have no money incomes to spend and no prices to pay, even the principles of demand, in so far as they concern prices, would fall by the wayside. The consumers in the system would be expected to experience diminishing utility or satisfactions as they helped themselves to increasing quantities of any particular good, and the existence of this tendency might influence society to turn to other lines of production before carrying the production of any given good too far, but the tendency could not, of course, have any effect on prices. Similarly, from the point of view of supply a communistic society would have to solve the problem of distributing its scarce means of production among their various alternative uses and would have to face the fact that the use of unduly large quantities of the productive agents in any one industry would serve to limit the quantities of other goods which could be secured. However no theory of price determination, as such, could exist.

Value Theory in Relation to Fascism. Since most industries and businesses were privately owned and operated under fascism, and since private enterprisers received money prices, paid money costs, and sought profits, it might have been thought that our theory of value had considerable validity for fascist economies. In actual practice, however, private enterprisers were told by the government what and how much to produce, how much to pay for materials, labor, and other things, how much they could have of various productive agents, how much they could charge for their finished commodities and services, and many other things. In such a situation, no theory of value developed for a capitalistic system would have any application. The only theory of price determination for fascism was that prices would be whatever the leaders of the government and ruling party wanted them to be.

QUESTIONS AND PROBLEMS

1. How does the demand for the product of an individual monopolistic competitor differ from the demand for the product of an individual competing seller or monopolist?

2. "The position of the monopolistic competitor, in trying to dispose of a fixed stock of a good in the short run, is decidedly inferior to that of the monopolist." Explain.

3. "In actual practice, the best decision of the monopolistic competitor as to the price at which to offer a perishable fixed stock in the short run may prove to be incorrect." Explain.

4. "The cost situation of the monopolistic competitor in the intermediate period is the same as that of the competitor or monopolist." Show whether you agree.

5. "The monopolistic competitor determines his output in the intermediate period in the same fashion as an individual enterpriser under any other condition of the market." Do you agree? Explain.

6. "The monopolistic competitor is sure to make a profit in the intermediate period." Show whether you agree.

7. "It would be somewhat more difficult for the monopolistic competitor than for the monopolist to decide whether it would be worth while to make extensive outlays for advertising." Explain.

8. "The expansion or contraction of a whole industry under monopolistic competition in the long run is likely to occur by means of increases or decreases in the size of the firms already in the industry." Discuss.

9. "An industry under monopolistic competition may operate under conditions of increasing, decreasing, or constant cost in the long run." Explain.

10. "In the long run, profits or losses tend to be eliminated in an industry operating under monopolistic competition and price tends to equal average cost of production." Explain.

11. "When monopolistic competition is crossed with oligopoly, the individual firms in the industry may make considerable profits even in the long run." Do you agree? Explain.

12. Summarize the theory of price determination under monopolistic competition.

13. "A considerable part of the theory of demand would still be valid in a socialized economy." Show whether you agree.

14. "It would be very difficult to determine the cost of producing an economic good under socialism." Explain.

15. "Since both the price and the cost of any good would be determined arbitrarily under socialism, the relationship between them would be of little significance." Do you agree? Explain.

16. "The planning authority under socialism would have to see to it that money income equaled money outgo over the system as a whole." Show whether you agree.

17. Discuss the probable validity of our theory of price determination under communism and fascism.

REFERENCES FOR FURTHER READING

Boulding, K. E., *Economic Analysis.* New York: Harper & Brothers, 1941, Chaps. 4–5, and 18–28.

Chamberlin, E. H., *Theory of Monopolistic Competition.* Cambridge: Harvard University Press, 1933, Chaps. 2, 4–5.

Clark, J. M., *The Economics of Overhead Costs.* Chicago: University of Chicago Press, 1923, Chaps. 4–6 and 23.

Davenport, H. J., *The Economics of Enterprise.* New York: The Macmillan Company, 1913, Chaps. 2–3, 5–8.

Henderson, H. D., *Supply and Demand.* New York: Harcourt, Brace and Company, 1922, Chaps. 1–5.

Marshall, Alfred, *Principles of Economics.* London: Macmillan & Company, Ltd., 8th ed., 1920, Book III, Chaps. 1–6; Book V, Chaps. 1–5.

Robinson, Joan, *The Economics of Imperfect Competition.* London: Macmillan & Company, Ltd., 1933, Chaps. 1–4.

Stigler, G. J., *The Theory of Competitive Price.* New York: The Macmillan Company, 1942, Chaps. 5–9.

XVII

Introduction to Distribution

In several of the earlier chapters of this volume, we noted that modern methods of production create tremendous quantities of the commodities and services which are desired by the individual members of society. However, the individual commodities and services created by roundabout, large-scale, and specialized methods of production are not consumed directly by the individuals responsible for their production, but are instead offered for sale on the market. In recent chapters we have therefore examined in some detail the process which determines the rates at which commodities and services are exchanged directly for money, and indirectly for each other.

Commodities and services are offered for sale on the market by business enterprisers, but these business enterprisers cannot retain for themselves all of the money income which results from the sale of the products. They retain a part which constitutes the remuneration for labor, land, and capital furnished in the productive process by the business enterprisers themselves, and another part which represents the recovery of money sums already paid out to other individuals who have provided land, capital, or labor for production. A third part must be paid out to individuals who have contributed to the productive process (through their land, capital, or labor) and have not yet been remunerated.

Fundamentally, then, it is correct to say that the money income which results from the sale of all commodities and services on the market is divided or apportioned among the individuals who have been responsible for the production of these commodities and services. The receivers of money income proceed to convert the money sums which they receive into desired commodities and services, and in this way the real income of society (consisting of commodities and services) becomes divided or apportioned among the individuals who have contributed to the productive process. This division or apportionment of the money and real income of society among individuals is referred to as the distribution of income, and it constitutes the third main division of our study of economic principles.

Approaches to the Study of Distribution. There are at least three points of view from which the study of the distribution of income might be

approached. In the first place we might study the way in which the total income of our economic system is divided into general shares, such as rent, interest, wages, and profits, and the factors affecting the relative sizes of these shares. Thus we might ask how it happened that the wages and salaries of American workers amounted to over 100 billion dollars in a certain year while interest and dividend payments came to only 10 billion dollars altogether, and what would cause the relative size of such total payments to increase or decrease. This approach might be called the study of distribution proper, but it is seldom attempted in works on the principles of economics, perhaps for the reason that there is comparatively little that can be said about it.

A second possible approach is to consider the factors which determine the rates at which the various agents of production are remunerated on a per-unit basis. Regardless of the total shares received by the owners of the different agents of production, we would seek, according to this point of view, a knowledge of the forces which determine the rates at which wages are paid to units of labor of various grades, the rates at which rent is paid for the use of units of land of various grades, the rates at which interest or rent is paid for the use of different kinds of capital goods, and the rate at which interest is paid for the use of savings or capital funds. The relationship between this approach and that described in the preceding paragraph is obvious. That is, rates of remuneration are convertible into total shares of income if we know the quantities of the various grades of productive agents that are employed at the various rates of remuneration. On the assumption that the various agents of production tend to be fully employed in the long run, the analysis of the factors which determine their various rates of remuneration is an important approach to the study of the distribution of income.

Finally, we might concern ourselves with the incomes received by individuals or families in our economic system, without paying any attention to the question of whether these incomes are received as rent, interest, wages, or profits, or some combination of these types of income. The study of distribution in terms of rent, interest, wages, and profits usually neglects the great disparities which are likely to exist in the amounts of income received by individuals and families. For some purposes at least, the important thing to consider is the amount of income received by the individual or family from all sources rather than the rates at which the individual or family receives remuneration for units of certain productive agents. We shall deal with the distribution of income on a personal and family basis to some extent in Chapter XXII, but the greater part of our discussion of distribution will deal with rates of remuneration for the productive agents.

The Theories of Value and Distribution. In studying distribution in terms of the rates at which the agents of production, or the various grades of the agents of production, are remunerated on a per-unit basis, it should be made clear that the theory of distribution is not something altogether separate and distinct from the theory of value. The theory of distribution may be viewed as an attempt to apply the principles of price determination to the prices paid for the use of the agents of production, and these prices, like those paid for finished goods, are determined by the forces of demand and supply. The costs of production of finished products consist fundamentally of the prices paid for the use of the agents of production, and it is the business enterprises producing and supplying the finished products which demand the agents of production.

In the theory of value, our interest in the prices of final products causes us to work from the prices of the productive agents to the prices of the final products, but in the theory of distribution our interest in the prices paid for the use of the productive agents causes us to proceed in the opposite direction. Actually, of course, it cannot be said either that the prices paid for the use of productive agents are causes of the prices of finished products or that the prices commanded by finished products are causes of the prices paid for the use of productive agents. Instead, the true situation is that the prices of finished products and the prices paid for the use of the productive agents are *simultaneously determined* in the market. With this in mind, it is seen that the theories of value and distribution look at essentially the same problem from different points of view and that the difference between them is one of emphasis.

THE DEMAND FOR THE AGENTS OF PRODUCTION

Business enterprisers demand units of any productive agent because of the contribution which these units are able to make to the output and value-product of the enterprises. Hence, in examining the basis of the demand for any productive agent, we must concern ourselves again with the analysis (originally presented in Chapter VI) of the productive results to be obtained by combining varying amounts of a given productive agent with fixed amounts of other productive agents.

Diminishing Productivity. Suppose that a business enterpriser is setting up a productive establishment in a certain line of manufacturing. He has his land under control for a long period of years, the factory is constructed, varied types of machinery, tools, and equipment have been purchased and installed, and adequate amounts of raw materials, supplies, and fuel are on hand. This enterprise could turn out some product with only a few workers, just so there were enough to generate the necessary power, operate at least one unit of each type of machinery and equipment part time, and perform other essential tasks. However, much of the power which was generated

would be wasted, much of the machinery and other facilities would be idle most of the time, and each worker would have to function on a nonspecialized basis. Both the total product of the enterprise and the output per worker would be rather small in this situation.

As more workers were used along with the fixed productive agents, the quantity of labor would become relatively more sufficient, better use could be made of the fixed productive facilities, and the workers could concentrate on individual specialized tasks. The result would be both an increasing total product and an increasing average product per unit of labor in this stage of operations. The enterpriser would scarcely wish to stop acquiring units of labor so long as additional workers would increase both the total product of the enterprise and the product per unit of labor. Sooner or later, however, the tendency to diminishing productivity per unit of the variable productive agent would assert itself. There are obvious limits on the specialization of labor by tasks and after a certain point had been reached additional workers would have to be assigned to less and less important tasks in the enterprise. In this stage of operations, each worker added would still be able to increase the total product of the enterprise to some extent but the average product per unit of labor would decline as additional units of labor were employed. It is somewhere in this stage of increasing total product but declining product per unit of labor that the enterpriser would want to stop taking on workers. In other words, the enterpriser would have no incentive to go beyond the point of absolutely diminishing returns, for beyond this point each worker could add nothing to the total product or would even cause the total product to decline.

Marginal Product and Marginal Productivity. In deciding how many workers to hire, the employer would be more concerned with the marginal product of labor and the marginal productivity of labor than with the average product per unit of labor. The marginal product of labor at any stage of operations is the amount of product which one more worker could add to the total product. The relationship between marginal product and average product is readily understandable. As long as the marginal product added by a given unit of labor is greater than the average product of all preceding units of labor, the average product per unit of labor must rise. In this connection it makes no difference whether the marginal product itself is rising or falling. Just so long as it is greater than average product, the average product must rise. On the other hand, when the marginal product added by a given unit of labor is smaller than the average product of all preceding units of labor, the average product of labor must fall. Since marginal product is greater than average product when average product is rising and smaller than average product when average product is falling, it follows that marginal product is equal to average product

when average product is neither rising nor falling; i.e., when average product is at its maximum.

It follows from this discussion that, if we plot curves of average and marginal product in a diagram, the marginal product curve rises more rapidly than the average product curve, reaches its peak at an earlier stage in the application of successive units of labor to the fixed productive agents, intersects the average product curve at the highest point of the latter curve, and then declines more rapidly than the average product curve. These relationships may be seen by a re-examination of Figure 1 on page 111 of this volume. The marginal productivity of labor is quite simply the money value of the marginal product of labor. That is, marginal productivity is marginal product multiplied by the price obtainable per unit of product. A curve of marginal productivity in a diagram would have exactly the same characteristics as a curve of marginal product, except that the marginal productivity curve would be drawn in terms of value-product rather than physical product.

Marginal Productivity and Wages. If we make the assumption that there is no established market rate of wages for the labor that he needs, the employer may stop hiring workers at any point in the range in which the total product of his enterprise is increasing but the average product per unit of labor is decreasing. Wherever the enterpriser stops taking on workers, the last worker hired may be considered for the moment to be the marginal unit of labor. The amount of product which he can add to the total product of the enterprise (or marginal product) when he works in conjunction with his proportionate share of the other agents of production, is smaller than that of any preceding worker. The money value of this marginal product, or marginal productivity, represents the largest sum which the enterpriser can afford to pay him as wages.

Moreover, a little reflection leads to the conclusion that this same marginal productivity represents the largest amount which the enterpriser can afford to pay for any other worker in the group. On the assumption that all the workers are of the same general grade or quality, they are taken to be interchangeable in the factory. If this is true, the withdrawal of any worker in the group will result in the loss of only the marginal product of the labor. If the first worker hired, or any other worker, withdraws from the enterprise, the last worker may be moved up to replace him and, in effect, the only worker lost is the last or marginal worker. Hence no worker in the group is more important to the enterpriser than the last or marginal worker and, in the situation now under discussion, all the workers will receive a wage in the long run which is equal to the value of the marginal product of labor, or marginal productivity.

This is, of course, the same thing as saying that any worker, out of a group of like workers, is last or marginal. The process of diminishing

productivity can be followed more easily if we assume that the individual workers are hired one after the other, but the conclusion as to wages would be the same if all the workers were hired at once. The last unit in the marginal sense is last in point of quantity and not in relation to time. In a group of 200 like workers, the last worker is the one who makes the number 200 instead of 199, and any worker in the group can be so regarded. Thus, if we disregard supply considerations for the moment, from the point of view of demand it is seen that the rate of remuneration which can be paid for an agent of production tends to equal marginal productivity, that marginal productivity is the money value of the marginal product, and that the marginal product is the amount of product which is attributable to the use of *any one unit* of a given supply of a productive agent. We have also developed a tool which will be of some use in connection with the analysis of the distribution of income. If we wish to find the marginal productivity of a given quantity of a productive agent, the proper procedure, at least in theory, is to withdraw one unit of the productive agent in question and observe the resultant decline which will occur in the total product turned out by the enterprise, or combination of productive agents.

The Demand for an Agent of Production. Returning to our illustration, we now assume that there is a prevailing rate of wages for labor of the sort which our enterpriser needs, and that he must pay it if he is to obtain any of the labor in question. In this situation it is not difficult to predict the general nature of the enterpriser's demand for this labor, if he is considered to be operating under competitive conditions. The amount that each worker adds to the total costs of the enterprise is his wage, while the amount that he adds to its total receipts is his marginal productivity (marginal physical product multiplied by the price per unit of the product). The enterpriser will go on hiring workers as long as each worker adds more to the receipts of the enterprise than to its costs, and he will cease hiring at the point where the marginal productivity of another worker would be inadequate to cover his wage. This simple conclusion is possible because, under competitive conditions, the individual enterpriser can change his rate of operation and output without affecting the price of his product or the cost per unit of any productive agent to him.

It follows that, if the wage rate is high for labor of the desired grade, the enterpriser will use a relatively small number of units of it because it is only by keeping this labor scarce in his enterprise in relation to the available quantities of the other agents that he can keep its marginal productivity high enough to justify the payment of this high wage rate per unit of labor. On the other hand, if the wage rate for this grade of labor is low, the enterpriser will employ large quantities of it in relation to the available quantities of the other agents. Its productivity will not be high under the circumstances, but it will not need to be high in order to cover

the low wage rate which has now been assumed. Thus the demand of the enterpriser for this labor will show that small quantities of it would be taken at high wage rates and large quantities at low wage rates, and the demand curve would move downward from left to right in a diagram like the demand curve for any economic good. Since the demands of other enterprisers for this labor would have the same general characteristics as that of our individual enterpriser, the total demand for this labor may be thought of as an aggregate or summation of the demands of the individual enterprisers and would also show that the quantity of this labor which all enterprisers would use would vary inversely with the wage rate.

If any other productive agent were substituted for labor in our illustration, the same conclusions would still be reached and the competitive enterpriser would be willing to go on adding units of the agent down to the point where the marginal productivity of another unit would be insufficient to cover its cost. In general terms, then, it may be said that the demand curve of a competitive enterprise for any productive agent is the same as a portion of the curve of marginal productivity for that agent in that enterprise. Since a competitive enterprise, as we have seen, always operates at some rate in the region in which the marginal productivity of any productive agent is declining, the demand curve of the enterprise for any productive agent also slopes downward from left to right in a diagram and indicates that the quantity which would be taken varies inversely with the cost per unit of the agent. The total demand curve for a productive agent has the same general characteristics as the demand curve of an individual enterpriser, since it is the summation of the demands of the individual enterprisers.

THE SUPPLY OF THE AGENTS OF PRODUCTION

Diminishing and marginal productivity are factors which operate on the demand side in determining the rates at which productive agents are remunerated, but these rates received by the owners of the agents are not determined by demand alone. Indeed, as we have noted, the marginal productivity of an agent of production, or grade of an agent, tends to vary inversely with the quantity of that agent which is available for use with given amounts of the other agents. It is necessary, therefore, in connection with the general theory of distribution, to consider the possible reflex influence which the rate of remuneration of an agent of production may have on the total supply of that agent.

The Supply of Land. Land has been defined as natural wealth or as wealth which exists without human labor having been expended to produce it. It is clear enough that the quantity of natural things in existence is independent of the prices which may be paid for their use, but there are at least two possible sources of variation in the effective supply of land.

For one thing, out of the total amount of land available, the quantity which owners will offer for use in short periods of time may vary to some extent with the size of the rent payments which are obtainable, since owners may hold some valuable land out of use temporarily if the annual rent payment which can be secured is unsatisfactory at the moment but is expected to improve later on. However, in the long run all of the soil and other natural things which are worth using may be expected to find their way into the effective supply of land on the market.

In the second place, the supply of land, as natural wealth, is capable of some variation in long periods of time because soil and other natural things, which are too poor in quality to be worth using at a particular time, may become worth using as time goes on. Thus soil areas or mineral deposits may be so poor in quality (or location) that they do not qualify as land or natural wealth at a given time because no one uses them and they do not command a price. Later on, when the population of the country and the total quantity of capital available for production have increased greatly, or improved methods of production have been developed, the same areas or deposits, without any change in their quality, may be used, command a price, and qualify as natural wealth and land. However, if soil areas have to be improved by irrigation or drainage before they can be used, they do not become "land," strictly speaking, even when they come into use, for they are really in part a product of labor and capital.

The relative fixity of the total amount of land or natural wealth should not lead one to the conclusion that the rents of land are set once and for all and cannot change as time goes on. Although no very large changes are to be expected in the total amount of land itself, the other agents of production which are used with it, labor and capital, are capable of great increases in total amount. As given amounts of land are used in conjunction with increasing amounts of labor and capital, the marginal productivity of labor and capital tends to decline (other things being equal) while that of the land tends to increase. Also, if other things do not remain unchanged and better machines and methods of utilizing land are developed, the marginal productivity of land, or the money value of its marginal product, may increase. Thus the rents of land, based on marginal productivity, may change from time to time even though the total quantity of land changes only slightly if at all.

The Supply of Labor. The total quantity of labor of a certain grade, or of all grades, if measured in terms of the number of workers in existence, is also fixed at any given time, but the number of workers who offer themselves for employment may vary to some extent with the size of the wage rates which may be obtained. Moreover, the hours per day or other period of time which the workers spend at their tasks and the intensity of the efforts which they put forth are capable of some variation. Hence the total

quantity of labor, if measured in terms of units of labor service, is by no means fixed even in short periods of time. While there is considerable controversy over the matter, perhaps the most common opinion is that the total quantity of labor in terms of units of labor service tends to vary directly with the wage rate in such periods of time. That is, more workers will come forward for employment, and those who come will work harder and longer, for high wage rates than for low ones.

In the long run, the number of individuals who offer themselves for employment, the length of the working day, and the intensity of effort on the part of the workers are regarded as comparatively constant factors, and the total quantity of labor is thought to depend rather directly on the growth or decline of population, or of the number of potential workers in existence. Many writers have attempted to trace a relationship of one kind or another between population growth, or changes in the number of potential workers, and wage rates, but these attempts have been rather unsuccessful. In Chapter XX we shall examine the reasons for thinking that such changes in numbers or population are largely independent of wage rates. For purposes of wage discussions, labor of the managerial or enterprising variety will be grouped with labor of other grades.

The Supply of Capital. In the short run both the total quantity of capital goods and the total quantity of fluid savings or capital funds available for investment in such capital goods may be taken to be fixed. In the long run all savings or capital funds must be thought of as fluid or capable of being invested in any desired sort of capital goods, and the total quantity of such savings tends to vary directly with the rate of interest. To be sure, some units of savings are saved for other reasons than to obtain interest, as we noted in Chapter IV, but there are many units of savings which would not be made if it were not for the payment of interest on such funds. Since there are many units of money income which their owners would rather spend for consumption now instead of later, an increase in the rate of interest is expected to lead eventually to an increased amount of postponement of consumption and a larger total quantity of savings. Conversely, a decline in the rate of interest is supposed to induce people to spend larger portions of their money income for present consumption and put aside smaller portions as savings. These effects on the total quantity of savings will tend to follow even though many persons save various amounts for other reasons than to obtain interest.

General Conclusions. While it is clear that variations in the available quantities of various productive agents, or grades of the agents, will have an effect on the rates per unit at which the owners of these agents are remunerated, such variations in quantity affect the rates of remuneration through their influence on the marginal productivity of the agents. Therefore, under any conditions of supply and under the assumed conditions

of distribution theory which will soon be explained, our previous conclusions, as stated in discussing the demand for the productive agents, hold good. That is, the owners of any agent of production, or grade of an agent, will be remunerated at a rate which, though determined by demand and supply, is equal to the marginal productivity of the agent. This marginal productivity is the money value of the marginal product, and the marginal product is the amount of product which is attributable to the use of any one unit of a given supply of a productive agent.

ASSUMPTIONS OF THE THEORY OF DISTRIBUTION

As we have just implied, our general conclusions with regard to the determination of the rates of remuneration of the productive agents, and the analysis on which these conclusions rest, are not advanced as being valid in all times, places, and conditions. Instead, they depend for their validity on a number of assumptions concerning conditions of the market for the productive agents and other matters, and we must now undertake an analysis of these assumed conditions.

Capitalistic Institutions. In the first place, we may note that the distribution of income is likely to vary considerably from one type of economic system to another. Since we are primarily interested in economic activities and practices under capitalism, we shall assume in the main body of our discussion of distribution that the various agents of production are privately owned, that their owners receive the earnings of the agents as private income, that the motives of the owners of the agents are primarily economic in character, and that the owners are free to direct their agents of production into such lines of economic activity as seem most appropriate on the basis of prospective earnings. Then, at the end of the discussion of distribution, some reference will be made to the probable characteristics of income distribution under other types of economic systems.

Real Income and Money Income. Our chief interest is also in the distribution of the *real income* of our capitalistic system among the owners of the agents of production, but it is usually more convenient to deal with money incomes than with real incomes in discussing questions of distribution. Hence we should like to be dealing with a situation in which changes in money incomes always represent like and proportionate changes in real incomes. This situation does not exist in practice because the general price level and purchasing power of money change from time to time. However, we can create the desired situation artificially for our present purposes by making in the study of distribution the same assumption which was made in the study of price determination; i.e., that for the purposes of this study the price level and purchasing power of money remain constant.

The Competitive Assumptions. The general theory of distribution also depends upon the existence of competitive conditions in the markets for

the agents of production. This means, in the first place, that the owners of an agent of production and those who demand it in production should have a fairly wide knowledge of the market conditions which concern that agent. The owners of the agent should know what its marginal productivity is in a given employment and should be aware of alternative employments which exist and the rates of remuneration available in these other uses. The enterprisers who demand the agent should know what it is worth to them and what is paid for it in other uses. Second, there must be many persons supplying the agent and many persons demanding it. The number of persons demanding an agent should be so large that it appears to the individual enterpriser that the agent is available to him in any amount at a constant supply price and that he can increase or decrease considerably the amount of the agent which he uses without affecting the price or rate of remuneration per unit which he has to pay for it. In similar fashion, the number of suppliers should be so large that any one owner can assume that he can offer more or less of the agent than usual without affecting the rate of remuneration which he will receive. Third, it is necessary that the users of the agents compete actively for available quantities of the agents and that the owners of the agents compete actively for the available opportunities for the employment of the agents. There must be no agreements or conspiracies or organizations on either side of the market. Any organized attempts on the part of the owners of the agents to force up or keep high the rates of remuneration of the agents and any organized attempts of enterprisers to force down or keep low these rates of remuneration constitute a violation of this assumed condition.

Fourth, our theory assumes that the various agents of production are mobile; that is, they are free to move from one employment to another or from one place to another and are capable of doing so if a comparison of rates of remuneration in various employments or places indicates that such transfers are desirable. Fifth, it is implied that there is substantial equality of bargaining power between the individual owners of the agents of production and the individual enterprisers who desire to use the agents, so that neither parties are able to take advantage of the others. Finally, the theory assumes that the government does not interfere with the process by means of which the rates of remuneration of the agents of production are determined and that these rates of remuneration are left entirely to the tender mercies of the forces of demand and supply in the market.

Much space could be used at this point in discussing the extent to which these assumed conditions actually prevail in various practical situations in the market, but it seems better to postpone most of this discussion until those chapters which discuss the determination of the rates of remuneration of the individual agents of production. For the present it may suffice to say that these various assumptions are seldom, if ever, completely

true or completely false as descriptions of actual market situations for the determination of rates of remuneration for the productive agents. The extent to which they are realized in practical situations varies from one time to another in regard to the same agent of production and varies from one agent of production to another at the same time. Thus, at a given time, the owners of land and capital may be rather well informed as to market conditions affecting their agents, while the owners of labor may be rather ignorant of market conditions in connection with wage determination, or the government may interfere significantly with wage determination, less with interest determination, and not at all with rent determination. On the other hand, the government may take an active part in the process of wage determination at one time and leave it entirely alone at another, or workers may be organized into unions to a much greater extent at one time than at another.

In some cases at least, the assumed conditions have a better chance to be realized in actual market situations in the long run than in short periods of time. For example, workers may not be able to move quickly from one occupation or place to another, but such movement may take place if disparities in wages and other conditions between occupations or places persist over longer periods of time. To the extent that such things are true, our conclusions and analysis should be regarded as more nearly valid in long periods than in short periods, and this is fortunate because our chief interest lies in the distribution of income in the long run. However, when the various assumptions are not realized in practice, or are only partly realized, the application of the marginal productivity analysis and the theory of distribution in general is weakened to that extent even in the long run. In general, it is wise to regard this theory of distribution merely as a rough general statement of tendency, but one which is perhaps somewhat more accurately descriptive of the distribution process than any other theory which can be devised.

The Influence of Monopoly and Monopolistic Competition. The marginal productivity analysis and the general theory of distribution which has been stated depend also on another assumption which is much more general than those which have just been given. This assumption concerns the existence of competitive conditions in the markets for the final products sold by the enterprisers who demand the agents of production, and is not the same thing as assuming competition among the enterprisers in demanding the agents of production. There could be a high degree of competition among enterprisers in demanding an agent of production even though the enterprisers sold their finished products under conditions of monopolistic competition.

The willingness of an enterpriser, under competitive conditions in connection with the sale of the finished product, to use an agent of pro-

duction down to the point where the last unit of the agent would cost a sum equal to the value of the additional product which the last unit of the agent could add to the total product of the enterprise is based upon the ability of the enterpriser to sell all of his product, both the original amount and that added by the last unit of the agent, at the price prevailing in the market. The enterpriser has this ability under competitive conditions because he furnishes such a small part of the total output of the good that he can increase or decrease his output without affecting the price per unit he can obtain for his product. However, under monopolistic competition or monopoly the individual enterpriser cannot sell as much or as little as he pleases of his product without affecting the price per unit which he can obtain. Instead, the price which he can obtain per unit of his product tends to vary inversely with the quantity of the good which he decides to sell. Consequently, he cannot hire an additional unit of an agent of production if the product which it will add, multiplied by the price per unit of the product, just equals the cost of the additional unit of the agent, because the presence in the market of the product of the additional unit of the agent will tend to lower the price at which the entire output can be sold.

In deciding to hire or not to hire an additional unit of the agent of production in question, the individual enterpriser under monopolistic competition or monopoly must follow a rather complicated analysis. He must multiply the units of product to be obtained from the additional unit of the agent by the price per unit which can be obtained for this product. He must multiply his previous units of output by the decline in the price per unit of the product which will occur if the product of the additional unit of the agent is thrown on the market along with his previous output. He must subtract the loss of revenue on the previous output from the gain in revenue which will result from the sale of the goods produced by the additional unit of the productive agent. The difference between these two sums must then be related to the cost of the additional unit of the agent, and the additional unit of the agent will be hired only if its cost is at least covered by this difference. Thus the marginal productivity of a productive agent must be computed by a monopolist or monopolistic competitor in a fashion somewhat different from that which is appropriate under competition, and the reason is that, whereas average revenue (price) and marginal revenue are the same for the competitive enterpriser, they are quite different for the monopolist or monopolistic competitor.

This discussion does not mean that competitive enterprisers will pay one price for the use of a certain productive agent while monopolists or monopolistic competitors will pay a different price for the use of the same agent. In the long run under competitive conditions in the market for an agent of production, all enterprisers will be expected to remunerate the

agent at the same rate per unit, whether they sell their finished products under conditions of competition, monopoly, or monopolistic competition. Neither does it mean that, if the rate of remuneration of the agent is equal to marginal productivity for competitive enterprisers, it will be higher or lower than marginal productivity in the case of monopolists or monopolistic competitors. What our discussion does mean is that, if many enterprisers who demand a certain agent of production sell their products under conditions of monopoly or monopolistic competition, the rate of remuneration (equal to marginal productivity) at which *all* of a given quantity of the agent can find employment will be lower than that at which the entire amount of the agent would find employment if all the enterprisers demanding it sold their finished products under competitive conditions. The effect of having a significant element of monopoly or monopolistic competition in the sale of finished products among the enterprisers who demand an agent of production is to make the total demand for the agent less elastic than it would be otherwise. That is, a given fall or rise in the rate of remuneration of the agent will not increase or decrease the quantity of it which is used as much as it would if competitive conditions prevailed in the market for the finished products.

Noncompetitive Conditions in the Markets for Productive Agents. If competitive conditions do not exist in the markets for the productive agents themselves, the reason is not usually that monopoly or near monopoly exists on either side of the market. Rather it is found in the fact that some one or more of the other conditions of competition fails to be realized in practice. However, cases of monopoly or near monopoly are possible, at least in theory, in the markets for productive agents. There is no danger of monopoly on the supply side for land, labor, or capital as a whole, or for large grades of land, labor, or capital, but if any agent of production is divided into very small grades, there is some chance that the entire amount of such a grade may be in the hands of a single person or firm, or a small number of persons or firms. If there is only one supplier of a certain small grade of an agent of production and no variable costs are involved in offering the agent on the market, the monopolist may be expected to try to maximize his total revenue derived from renting out his agent on the market. He will offer all of it for use if he can get the greatest total revenue in this way but he will not hesitate to hold a part of it out of use if he can get more revenue by renting out only a part of the amount which he owns than by letting it all be used.

If variable costs are involved in furnishing the small grade of the agent on the market, the monopolist may be expected to offer such a quantity of the agent that the marginal revenue which it produces will be equal to the marginal cost of bringing it to the market. He will not continue to offer it as long as the rate of remuneration or average revenue per unit

will cover the cost of furnishing the agent on the market, although this is what we should expect an owner of an agent of production to do under competitive conditions. Conditions of monopolistic competition are not very likely to exist in the supplying of an agent of production, because the different grades of an agent of production, where different grades exist, are likely to be the result of natural conditions rather than of conscious attempts at product differentiation on the part of the owners. However, a small grade of an agent of production may all be in the hands of a few suppliers or owners (oligopolists) and in such cases the few owners may be expected to behave as nearly as possible like pure monopolists.

Again there is little chance that land, labor, or capital as a whole, or large grades of land, labor, or capital, will be demanded by only one enterpriser or firm, or even a small number of enterprisers or firms. A very small grade of one of these agents of production might be so demanded. The situation would be called monopsony if only one enterprise demanded the agent and oligopsony if there were only a few enterprisers or firms demanding it. The theory concerning such cases is rather complicated in relation to its probable value, but it may be said in general that under these conditions, in the hope of obtaining it at a relatively low price per unit, the demander or demanders would probably restrict the use of the agent, in comparison with what would happen under full competitive conditions.

QUESTIONS AND PROBLEMS

1. "The study of the distribution of income might use any one of three approaches." Explain.

2. "The theory of distribution is merely the theory of value viewed from another angle." Show whether you agree.

3. "The analysis of the productive results to be obtained by combining varying amounts of a given productive agent with fixed amounts of other productive agents is of great importance in studying the demand for any productive agent." Do you agree? Explain.

4. "The rate at which the owners of any agent of production are remunerated for its use is determined by the marginal productivity of the agent." Do you agree? Explain.

5. Using capital as the variable agent of production, explain diminishing and marginal productivity.

6. "All like units of an agent of production will be remunerated in the long run under competitive conditions at a rate equal to the money value of the amount of product contributed by the last unit of the agent." Show whether you agree.

7. "The demand curve for an agent of production, or any one grade of an agent, has the same general characteristics as the demand curve for a finished economic good." Discuss.

8. An enterpriser says: "I play no part in determining the wage rate paid to my workers, for I must pay the going rate of wages and can decide only how many workers to employ at that rate." Discuss his opinion on this point.

9. Distinguish between average product, marginal product, and marginal productivity.

10. "The marginal unit of a productive agent is the last unit of that agent employed." Discuss.

11. "The supply of land, as the economist defines it, is completely fixed both in the short run and in the long run." Do you agree? Explain.

12. "While the quantity of each grade of land is fixed in the long run, the rents which are paid for the use of the various grades are highly variable." Show whether you agree.

13. "The quantity of labor of a certain grade, as measured in units of labor service, is variable even in the short run." Explain.

14. "The long-run effects of high or low wages on the total quantity of labor are unpredictable." Do you agree? Explain.

15. "Capital is the agent of production whose quantity is most responsive to price relationships in the long run." Show whether you agree.

16. Summarize the various assumptions of the theory of distribution.

17. "A consideration of the assumptions of distribution theory leads inevitably to the conclusion that the theory will be more nearly valid in the short run than in the long run." Do you agree? Explain.

18. "The general theory of distribution must be modified to some extent if it is to fit situations in which the enterprisers, who demand the productive agents, sell their products under conditions of monopoly or monopolistic competition." Show whether you agree.

19. "A monopolist or a monopolistic competitor cannot afford to hire an agent of production at a rate equal to marginal product times the price obtainable for the marginal product." Explain.

20. Is it likely that there will be a monopoly of demand or supply in connection with the agents of production themselves? Explain.

See References for Further Reading at the end of Chapter XXII.

XVIII

Rent

In turning to a more detailed discussion of rates of remuneration per unit for individual agents of production, or grades of such agents, we shall begin with rent. We mean by rent *the payment which is made (per year) for the use of land,* as distinct from payments which are made for the use of either capital goods or capital funds (savings). Rent is received by the owner of a piece of land of good quality, whether he uses the land himself or lets it out to a tenant. If he uses the land himself, the rent comes in as part of the income of his enterprise and is called *implicit rent,* while a payment received from a tenant for the use of land is known as *explicit rent.*

Land and Land Improvements. In our discussion of rent we shall concentrate attention for the most part on land as soil or site areas and not on the other natural things which are grouped under the heading of land. Moreover, land in this discussion means the bare land and does not include any improvements which have been made on it. Houses, barns, other buildings, fences, drains, and other improvements are in the nature of capital goods and, while the return from their use may seem almost like rent in the short run, it tends to be common interest in the long run and should not be included in the rent of land. Rent means, therefore, the annual payments which are made for the use of the land itself, as if it were unimproved. Actual money payments from tenants to landlords may include both rent and interest and are sometimes known as *contract rent.* Payments for the use of the bare land are called *economic rent.* Clearly it is possible for a money payment or contract rent from tenant to landlord to be composed entirely of interest on improvements and thus to contain no element of economic rent.

The Use of Land. Rent is a payment for the *use* of land and not for *using up* the land. That is, it is assumed that the productive qualities of the land are maintained intact throughout the uses for which the rent payments are made. Site land may be used indefinitely without in any way destroying or even diminishing the usefulness of the land for this purpose.

In similar fashion, while agricultural land may be allowed to run down through improper use, its fertility is not strictly irreplaceable and can be kept up indefinitely through the use of familiar methods and practices. In other cases, as in the case of mining land, for example, the use of the land inevitably involves using it up. That is, the use of mining land involves the gradual removal of a scarce and irreplaceable natural resource and the user of the land is in effect buying the mineral deposit gradually. The owner must be compensated for the depreciation of his mineral deposit as well as for the use of the land. In such cases, the price which is paid for the use of the land is called a royalty rather than a rent and it is usually paid as a price per unit of the mineral removed. It will be so much per ton of coal or ore, or per barrel of petroleum, and not so much per acre. The amount of the royalty per unit of the mineral will ordinarily depend on the quality of the mineral obtained from the land and the ease or difficulty with which it can be removed from the ground.

Use Price and Purchase Price. We should also remember that rent is a payment for the use of land, with ownership remaining unchanged, and not a payment (purchase price) which is made in transferring the ownership of land from one person or firm to another. However, while it is not at all the same thing, the payment or price necessary to buy a unit of land is closely related to the payment which is made to obtain its use in production. Under competitive conditions, the market values or prices of pieces of land tend to be bid upward or downward through time to such an extent that, however large or small may be the return derived from the use of those pieces of land, it will only amount to common interest on the market values or prices.

Thus if a piece of land produces an annual income of $1,000 (rent) and the rate of interest which can be obtained for savings placed in any equally safe investment amounts to 5 per cent, it is common to assume that the $1,000 annual rent from the land should be considered as 5 per cent of its value. Accordingly the market value of the piece of land will be bid up to about $20,000 or twenty times its annual income and, other things being equal, this is the price which can be obtained if the land is to be sold. This method of figuring the value of a piece of land by using the annual rent payment and the going rate of interest is known as the capitalization process. It can be applied to fixed capital goods as well, if we bear in mind the fact that the useful life of such capital goods is short in comparison with that of land.

On the basis of the capitalization process, the value or purchase price of a piece of land which yields a given annual rent varies inversely with the prevailing rate of interest. If the going rate of interest declined from 5 per cent to 2 per cent, the $1,000 annual rent furnished by a piece of land would be considered equal to 2 per cent of its value rather than 5 per

cent, and the value or purchase price of the piece of land would be $50,000 rather than $20,000. On the other hand, if the going rate of interest increased from 4 per cent to 6 per cent, the annual rent would be considered equal to 6 per cent of the value of the land rather than 4 per cent, and the value or purchase price of the land would fall from $25,000 to $16,667.

Why Rent Is Paid. Rent is paid for the use of certain grades of land in our economic system because this land is able to assist in the processes of production and because it is scarce or limited in quantity in relation to the demand for it. The amount of rent which will be paid annually for the use of a given piece of land depends on demand and supply factors, such as the quality, fertility, or location of the land, the economic goods which can be produced upon it, the prices which can be obtained for these economic goods, and the available quantity of land of this grade. All things considered, the fertility of agricultural land is probably more important for purposes of rent determination than its location, although location may also be a factor of considerable significance. In the case of land which is used for site purposes, the matter of location is the all-important factor. Rent is received as *private* income in our economic system because, under capitalism, land is, for the most part, owned and controlled by private individuals and firms. In other economic systems land might be publicly owned and no charge would necessarily be made for its use.

Complicating Factors in the Study of Rent Determination. The study of rent determination would be greatly simplified if either of two things were true of the land available for use. If there were several grades of land each of which could be used for only a single economic purpose or line of production, or if there were only one grade of land which could be used for several economic purposes or types of production, we should have little difficulty in presenting the discussion. Actually, however, there are several possible productive uses for each grade of land and there are several different grades of land which can be used in each general line of production. Since there is no very effective way of considering rent determination from both of these points of view at the same time, we shall consider the two situations in order, starting with rent determination for a single grade of land which is used in several different lines of production, and, by referring from one to the other, we shall establish the connection between the two situations. The situation which we are to study first can be handled by the ordinary demand and supply analysis, because it deals only with a single grade of an agent of production. The other situation involves several grades of land and must be dealt with by a special type of analysis.

RENT DETERMINATION FOR A SINGLE GRADE OF LAND

Demand and Supply Analysis. In considering rent determination for a single grade of land which is used in several lines of production, let us

take as our illustration first-grade land for site purposes in the downtown area of a small city. The exact amount of land in this downtown area which might be considered superior for site purposes would be difficult to determine in practice, but we shall assume in this case that it includes all land which lies in any direction within three blocks of the city hall. Clearly the amount of land in this arbitrarily assumed area is fixed. It is demanded for several different types of productive activity. Some of it will be used as sites for retail stores; other parts as locations for hotels, theaters, office buildings, and so on. In fact, some of the land might conceivably be demanded for consumption purposes as sites for homes, but we shall simplify our analysis by assuming that this condition is not present in the situation under discussion.

The characteristics of the demand for this land in each of the various lines of production are well known. A moderate quantity of this land in a given use, when combined with considerable amounts of the other agents of production, would be highly productive. However, beyond a certain point, more and more land of the same quality, used with fixed amounts of the other agents of production, would have less and less productivity per unit although it would continue to make some contribution to total product. In all probability the marginal productivity of the land in any one use would decline so rapidly that it would be economically impossible to use all the land in a single line of production at any positive rate of rent per year for each unit of land. It follows that each line of production would demand some of the land even at a very high rent, larger amounts of it at a moderate rent, and still larger amounts at a low rent, but any one line of production would not use all the land at any rent per unit which might prevail.

The Determination of Rent. The results to be expected under these conditions of demand and supply, in the long run under competitive conditions, are shown in Figure 32. The line S indicates the fixed amount of this first-grade site land under the assumed conditions and shows that all of this grade of land would tend to be offered for productive use, in the long run under competitive conditions, at almost any annual rate of rent which might prevail. The demand situation is that which was described in Chapter XIV as one of composite demand. The demand curve d^1 shows that a few units of the land would be used as theater sites even at a very high annual rent per unit, a larger number of units would be used for this purpose at a moderate annual rent per unit, and a still larger number of units at a low annual rent per unit. The demand curve d^2 shows the combined demands for this land as sites for theaters and sites for hotels. In plotting this second demand curve, to the units of land which would be taken for use as theater sites at the various annual rents per unit we have added those which would be taken as sites for hotels. In other words,

the various quantities which would be demanded as hotel sites have been plotted to the right of the d^1 line with that line as a base instead of the OY axis as the base. In similar fashion, the demand curve d^3 shows the demand for the land for use as sites for theaters, hotels, and office buildings, while the demand curve D shows the total demand for this land for all site purposes.

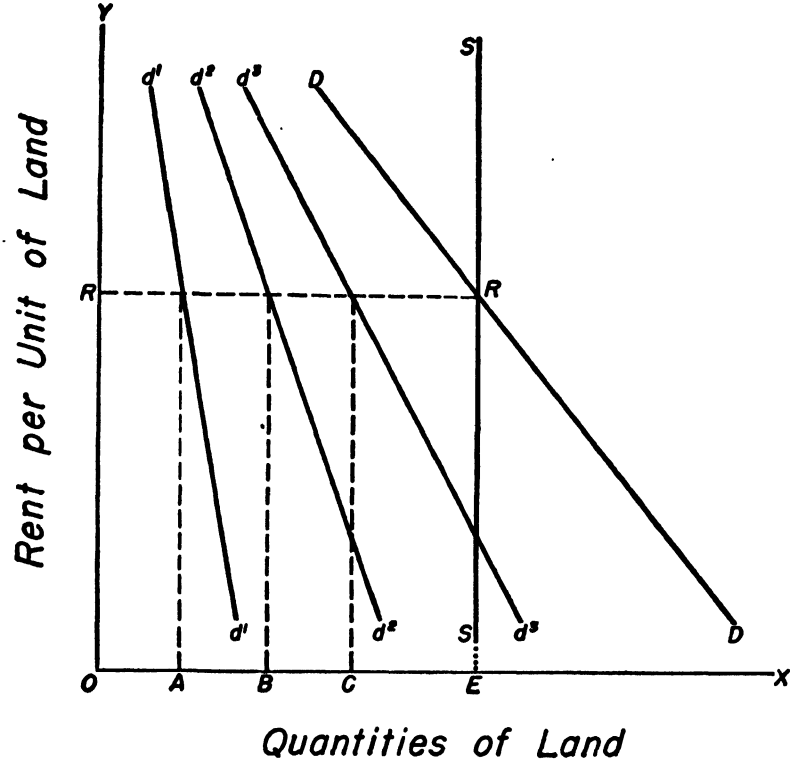

FIGURE 32.—Rent Determination for a Single Grade of Land with Several Uses in the Long Run under Competitive Conditions

The annual rate of rent per unit for this land in the long run under competitive conditions is determined at the point where the total demand curve D cuts the supply curve S. A broken line R is then extended from this point of equilibrium to the OY axis, cutting each of the lines representing partial demand. Further broken lines, dropped from the points of intersection of the line R with the partial demand lines to the OX axis, indicate the way in which the total available quantity of this land would be divided among the various individual uses or lines of production. That is, assuming that there are only four uses and demands for the land, OA

units of the land will be used for theater sites, *AB* units for hotel sites, *BC* units for office building sites, and *CE* units for retail store sites.

Some of our conclusions with regard to rent determination for a single grade of land would almost certainly not be valid *in the short run*. For example, Figure 32 shows all of the land of this grade to be in use for some productive purpose or other. This might not be true at any particular time but it will tend to be true in the long run under competitive conditions. In similar fashion, the diagram shows that all enterprisers in various lines of production pay the same annual rent per unit for land of this sort. It would be amazing if this were true at any given time, but it tends to be true in the long run under competitive conditions, since landowners are in general economically motivated and will not turn their land over to one productive use if a higher rate of rent can be obtained with equal safety in some other use. At any one time we should expect to find different pieces of the land rented out for varying periods of time under leases contracted at different times under different conditions of general business, and hence at different rates of rent per unit. In the long run under competition we should expect to find the total amount of this land to be distributed among the various lines of production in such a way that it commanded the same rate of rent per unit in each of its several alternative uses.

The Probable Validity of Assumptions. Now let us consider the extent to which the assumptions of distribution theory are likely to be realized in practice in the market for this land in the long run. First, both the owners and tenants of this land are likely to be fairly well informed as to the marginal productivity of the land and the market conditions concerning it. Second, the chances are good that there will be many people demanding the land and many people supplying it. It would be possible for all of the land in the area under discussion to be owned by one person or by a very few persons, but such a situation is not likely. Again it is rather probable that the owners of the land will compete actively in the matter of obtaining tenants and that prospective tenants will compete actively for the use of the available units of land. It is not especially likely that there will be agreements or conspiracies among either landlords or tenants for the purpose of controlling rents, although it is sometimes alleged in particular local areas that landlords have banded together and conspired to push rents up or keep them high.

The mobility of land of a certain grade between occupations or industries is likely to be very limited at any particular time because individual pieces of land are often tied up for long periods of time in some line of production or other under leases which prevent the landlords from transferring their land to other uses which might pay higher rents. However, not all land is tied up in this way, and in any case in the long run there

is nothing to prevent the transfer of land from one industry to another. With regard to the downtown urban site land which is under discussion, bargaining power will probably be substantially equal between landowners and tenants. In other cases, however, the bargaining power of the landowners may be much greater than that of the tenants. Finally, the probability that the government will interfere significantly with the market process by means of which rates of rent are determined does not seem very great, except in wartime or under other exceptional conditions. All things considered, the various assumptions of the theory of distribution probably have as good a chance of being realized in practice in connection with rent determination as in any other part of the field of distribution, and are realized to a greater extent for rent determination than for the determination of wages.

Thus our conclusions as to rent determination probably amount to a fairly accurate description of what will happen in actual markets for land in the long run, provided that the enterprisers who demand the land sell their products under competitive conditions. If the users of the land sell their products under conditions of monopoly or monopolistic competition, our analysis requires some modification. The supply situation would not be affected, and the various individual demand curves would still move downward from left to right in the diagram, but they would be less elastic than those shown in Figure 32. That is, greater declines in annual rents would be required to induce enterprisers to take larger quantities of the land, since the rent that could be paid for the use of an additional unit of land would not equal marginal productivity as computed under competition but rather marginal productivity as discounted by the reduction in price per unit which the enterpriser would have to take on his previous volume of output. In the case of land used for agricultural purposes, the chances are good that the enterprisers who use the land will sell their products under conditions that closely approximate those of competition, but many of the enterprisers who use land for site purposes are likely to be selling their products under conditions other than those of competition.

Land of Other Grades. Of course, the enterprisers who wish to operate theaters, hotels, office buildings, or retail stores in the fair city with which we have been dealing do not have to use first-grade land in the downtown area as sites for their businesses unless they want to do so. The theater operator or other enterpriser can locate his theater or business a few blocks away from the main business section of the city and make use of second- or third-grade site land (unless prevented by zoning restrictions), or if he wishes he can even locate his enterprise on the outskirts of town on low-grade site land. In general, however, there is little incentive to take such action, for the decline in the volume of business which can be done on low-

grade sites as compared with high-grade sites should be, in the long run under competition, just about large enough to offset the difference in the rents which have to be paid for the use of the sites.

RENT DETERMINATION FOR SEVERAL GRADES OF LAND

The Nature of the Situation. We now turn to a discussion of rent determination for the several grades of land which may be used for a given economic purpose, such as corn raising, for example. The land which is used for raising corn is not all of the same grade or quality. Just how many grades of land there are for this purpose would be difficult to determine in practice. Since our methods of analysis and conclusions do not depend upon knowing the exact number of grades of this land, it will be assumed arbitrarily that there are four such grades, as shown in Figure 33. In this illustration, the letter A represents one unit of first-grade cornland, the letter B one unit of second-grade cornland, and so on. In all cases it is assumed that there are many other units of each grade of land besides the unit which appears in the diagram.

Each rectangle which appears above a letter in Figure 33 represents one unit of labor and capital applied to the unit of land in question, and the product which results therefrom. Units of land of various grades are cultivated with varying numbers of units of labor and capital, it being assumed that the units of labor and capital are large enough so that diminishing productivity sets in on each grade of land after the application of one of these units. The first installment of labor and capital applied to the unit of Grade A land results in the production of 32 bushels of corn, a second unit of labor and capital adds 24 bushels of corn, a third unit adds 16 bushels, and a fourth unit 8 bushels. Grade B land is of poorer quality than Grade A land, and the first unit of labor and capital applied on a unit of it results in the production of only 24 bushels of corn, a second unit of labor and capital adds 16 bushels, and a third unit adds 8 bushels. On Grade C land, the first installment of labor and capital leads to the production of only 16 bushels of corn, and a second unit adds 8 bushels. Finally we come to Grade D land, which is so poor that the first unit of labor and capital applied to it results in the production of only 8 bushels of corn. Before the student begins to think that it is wonderful beyond all understanding that Nature should have decreed that the productivity of labor and capital on the various grades of cornland must always decline by 8 bushels at a time, it should be said that the figures have been arranged in this way arbitrarily in order to achieve simplicity in presentation. The *method* of computing rent is the same whether the differences in productiveness are taken to be uniform or varying, and we can pass later on to somewhat more realistic conditions.

The Marginal Product of Labor and Capital. Although different total quantities of labor and capital are applied to units of land of different grades, it should be noted that the "last" unit of labor and capital applied to a unit of each grade of land furnishes the same amount of added product. This amount of product (8 bushels) is the marginal product of labor and capital in corn raising. It is not the marginal product of the land, for that is one thing which we are trying to determine by means of this illustration. The application of the fourth unit of labor and capital to the unit of Grade A land is made because 8 bushels of corn are apparently enough to pay for one unit of labor and capital. If this were not true, the fourth unit of labor and capital would not be applied to this land. Now it follows that, if 8 bushels of corn are enough to pay for a unit of labor and capital

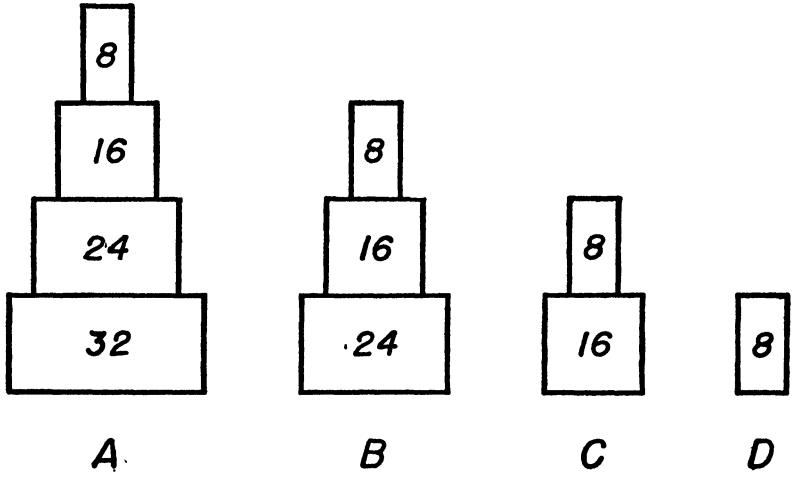

FIGURE 33.—Productive Relationships among Four Grades of Cornland

on Grade A land, the same added quantity of corn will pay for a unit of labor and capital on any grade of land. For this reason it is to be expected that the cultivation of each grade of land will be carried if possible to the point where one unit of labor and capital adds only 8 bushels of corn. Of course the third unit of labor and capital applied to Grade B land might add only 6 bushels of corn instead of 8. In that case, if the units of labor and capital could not be divided in any way, the cultivation of this grade of land would have to stop with the use of 2 units of labor and capital per unit of the land, because in the long run under competition the cultivators of this grade of land would have to pay as much per unit of labor and capital as the cultivators of any other grade of land.

Marginal or No-Rent Land. Grade D land is so poor for purposes of corn raising that the first unit of labor and capital applied to a unit of the

land results in the production of only 8 bushels of corn. Since this amount of product is just enough to pay for a unit of labor and capital, it follows that the users of this grade of land cannot afford to pay anything for the use of the land itself. If they were required to pay rent for the use of the land, they could not pay for the labor and capital which they employed, and this grade of land would not be used in the long run. Consequently Grade D land is called marginal or no-rent land. Students are sometimes puzzled by the concept of no-rent land because, so they say, they do not know of any land that can be had for nothing. It is necessary to remember, however, that the fact that no economic rent is paid for the use of a piece of land does not mean that no payment of money passes from the tenant to the landowner in connection with it. Such a money payment may exist even for no-rent land, but since it is composed of other returns, such as payments for the use of buildings or other improvements which the landowner has made on the land, it contains no element of economic rent.

Rent as a Differential Surplus. There are several ways of computing the economic rent of the better grades of land in our illustration. Needless to say, all methods produce the same answer with respect to the amount of economic rent on land of a given grade. For our purposes only two methods of computation will be explained. The first of these is the traditional method which explains rent as a differential surplus. Suppose that we want to find the economic rent of Grade A land in our illustration. If a unit of Grade A land is withdrawn from circulation, the gross loss of product which will result amounts to 32 plus 24 plus 16 plus 8, or a total of 80 bushels of corn. But these 80 bushels of corn are not produced by the unit of Grade A land alone, but rather by the unit of land in conjunction with four units of labor and capital.

If a unit of Grade A land is withdrawn from cultivation, the four units of labor and capital formerly employed on it will be set free for use elsewhere, and the loss of 80 bushels of corn will be reduced by the amount of product which these four units of labor and capital can produce in this other use. Since all the better grades of land are likely to be in use, this method of analysis supposes that the four units of labor and capital will be used on marginal or no-rent land. They will not be applied to one unit of marginal land, for one unit of labor and capital, employed on one unit of this land, produces just enough to pay for its cost. When the four units of labor and capital are applied, as a result, to four units of marginal land, they will produce a total of 32 bushels of corn. The 80 bushels of corn originally lost minus the 32 bushels of corn recovered gives 48 bushels of corn as the net amount of product which would be lost as the result of the withdrawal of one unit of Grade A land from cultivation. This is the marginal product of Grade A land and its money value (if corn is taken to be worth $1 per bushel) is $48, which is the marginal productivity of

Grade A land. In the long run under competitive conditions the economic rent of this land, as determined by demand and supply conditions, will tend to be $48 per unit, if everything remains unchanged.

Thus according to this analysis economic rent tends to be equal to the money value of the difference in total product between what a given quantity of labor and capital can produce on a certain grade of land and what the same quantity of labor and capital can produce when applied at the extensive margin of cultivation (that is, when appropriately applied to marginal or no-rent land). The marginal product, marginal productivity, and economic rent of the other grades of land in our illustration can be computed by the same method.

Another Approach to Rent Determination. The other method of rent computation deals with only one grade of land at a time. Considering Grade A land again we see that one unit of this land, when combined with four units of labor and capital, achieves a total product of 80 bushels of corn. The problem is to divide this total product into two parts, the one attributable to the land and the other attributable to the labor and capital. On this land the marginal product of labor and capital, or the product of labor and capital at the intensive margin, is 8 bushels of corn. We know from the analysis of marginal productivity that no unit of labor and capital can be worth more than the marginal product of labor and capital; hence no more than 32 bushels of corn, or 8 bushels per unit of labor and capital, can be attributed to the four units of labor and capital which are used on a unit of Grade A land. The 80 bushels of corn produced by land, labor, and capital minus the 32 bushels which are attributable to the labor and capital leaves 48 bushels which are attributable to the land itself. This, again, is the marginal product of Grade A land, its value (say $48) is the marginal productivity of this land, and in the long run under competitive conditions the annual economic rent per unit of this land tends to equal this sum. In similar fashion, under the assumed conditions, the annual economic rent of Grade B land could be computed as $24 per unit, that of Grade C land as $8, and that of Grade D land as zero.

The Influence of Increased Quantities of Labor and Capital. It is possible that under certain conditions there may not be any no-rent land. If the quantity of labor and capital available for corn raising increases, another unit of labor and capital may be applied on each unit of each grade of cornland. Such an additional unit of labor and capital might add another (say) 4 bushels to the total product on each unit of each grade of land. If the price of corn remains unchanged, the annual rent of each grade of land will go up. Under the second method of computation, the rent of Grade A land will become $84 minus $20, or $64 instead of $48 per unit. That of Grade B land will increase from $24 to $36 ($52 minus $16), that of Grade C land from $8 to $16 ($28 minus $12), and that of Grade

D land from zero to $4 ($12 minus $8). Thus even Grade D land, formerly marginal, will command a rent. Now, of course, if there is a fifth grade of cornland which was formerly not used because its quality is so poor that the first unit of labor and capital applied to it will bring in only 4 bushels of corn, under the new conditions this land will tend to come into use and will be the new marginal or no-rent land. If no such fifth grade of cornland exists, or if it is so poor in quality that the first unit of labor and capital applied to a unit of it would produce only 1 to 3 bushels of corn, all the grades of land in use for corn raising may command economic rent. In either case, the *tendency is for economic rents to increase in amount on the various grades of land, other things being equal, as labor and capital for use in conjunction with the land become more plentiful.*

Uneven Declines in Marginal Product. In our illustration of rent determination for the various grades of land used in corn raising, it was noted that the marginal product of labor and capital, on each grade of land above the marginal or no-rent land, was assumed to decline by 8 units of product as each additional unit of labor and capital was applied. It is now time to show that these marginal products could have declined at a varying rate without affecting the principles of rent determination. For example, on Grade A land the successive amounts of product could have been (say) 25, 18, 12, and 8 bushels of corn instead of 32, 24, 16, and 8 bushels. In this case the marginal product of the land would have been 31 bushels instead of 48 bushels, but it would have been determined by the same method as before; that is, by adding these successive products to get the total product and by subtracting from this total the product derived by multiplying the marginal product of labor and capital by the number of units of labor and capital applied to one unit of land. If it takes 8 bushels of corn to pay for a unit of labor and capital, each grade of land will be cultivated to the point where the last unit of labor and capital applied adds precisely 8 bushels of corn to the total product, if this is possible. Otherwise, if the units of labor and capital cannot be divided, the cultivation will tend to stop with the last unit of labor and capital which adds more than 8 bushels of corn to the total product. If the successive products on Grade B land are 22, 15, and 6 bushels of corn instead of 24, 16, and 8, only two units of labor and capital will be applied to each unit of land if an added product of 8 bushels is necessary to pay for a unit of labor and capital and a unit of labor and capital cannot be divided.

OTHER CONSIDERATIONS

Rent and Diminishing Productivity. The discussions of rent determination have suggested the close connection which prevails between the existence of economic rents for the use of land and the tendency toward diminishing productivity in the use of land. When the best grade of land

for a certain purpose is all in use and is not undercultivated, the use of additional units of labor and capital in that line of production, under given methods, tends to result in diminished added amounts of physical product, whether new land of poorer quality is brought into use or the best land is cultivated more intensively. Under these conditions, rent tends to appear, for it becomes worth while to pay something for the use of the best grade of land. Rent would not exist if the best land were unlimited in quantity, since no one would pay anything for the use of a free good. Similarly, rent would not exist if the Law of Diminishing Productivity did not apply to the use of a fixed amount of land with increasing quantities of labor and capital, for in this case the quantity of the best land would be, in effect, unlimited even though there was actually a fixed amount of it. To put the matter another way, if each unit of labor and capital applied on a unit of the best land for corn raising always added another 25 bushels of corn to the total product, the marginal product of labor and capital (25 bushels) multiplied by the number of units of labor and capital applied would always exhaust the total product and leave nothing for the marginal product of the land.

Is Rent an Earned Income? The question whether rent is or is not an earned income has aroused much controversy. The answer seems to depend somewhat on whether one takes the point of view of society as a whole or that of the private individual. From the social point of view it is not difficult to argue that rent is an unearned income and that the private landowner does not deserve to receive it. Although we may admit at the outset that the landowner deserves to receive a return from any improvements which he makes on his land, it may be pointed out that the landowner is not responsible for the existence of his land, for its natural fertility, or for its location. Land, as such, is a free gift of Nature and in an early stage of society is also likely to be a free good. Under the institution of private property it is later appropriated and owned by private persons and firms. As time goes on, the economic system develops and labor and capital increase in amount, land becomes scarcer in relation to these agents of production and in relation to the uses which people would like to make of it, and a price or rent can be secured for its use.

Not only does rent appear but its amount per unit of land tends to increase as economic activity expands and capital and labor continue to increase in quantity. A similar increase in the value or selling price of land occurs. The owners of land may be very energetic and industrious citizens, but in so far as economic rent is concerned it really makes no difference whether they are or not. By the very fact of landownership, all they have to do is sit and wait until larger and larger rents drop into their laps. Such rents, it is argued, are produced by social growth and development and not by any activity of the landowners. They are an unearned increment re-

ceived by these owners and rightfully belong to society as a whole. To be sure, the landowners may have borne risks or have exercised foresight in connection with landownership, but these activities should be rewarded only where they produce some social benefit and it is argued that the land would have increased in value just as much in the long run even if no private person had ever borne risks in connection with landownership, or speculated in land, or exercised foresight with regard to it. Since rents result from factors for which landowners are not responsible, it is argued that economic rent should be abolished as a type of private income and should be appropriated by the society as a whole which created it. This analysis of the nature of economic rent is the basis for what is called the single-tax program or movement.

Of course the argument that economic rent is an unearned income is not allowed to go entirely unanswered. While we may admit that many landowners received an unearned increment from landownership in the past, it is held that in many cases these owners have managed to escape with their booty by selling their land. Many landowners today have purchased their land at a price determined by capitalizing the actual or prospective annual economic rent at the going rate of interest and now receive a payment of rent annually which amounts only to common interest on their investment in the land. Why, it is asked, should a person who has saved funds and invested them in land be deprived of all return on his investment while another person who has saved a similar amount and made his investment in capital goods is allowed to receive a similar amount of income without interference? Indeed, if we wanted to tap the unearned increment from land it would be necessary to take interest from some individuals. That is, some landowners have sold their land at a high price and have invested the proceeds in other assets so that they are now inconspicuous interest-receivers. The increases in rent which they originally received were capitalized into the selling price of their land and, now that they have changed their investments, merely by taking interest they continue to receive what is in effect an unearned increment. Moreover, it is contended that even the proposal to tax away future increases in economic rents would work injustices in practice, since pieces of land are often sold at prices which represent the capitalization of prospective increases in economic rent as well as that of the existing rent payments.

Many years ago the question whether private landowners should be allowed to receive rent was almost never discussed, although interest as a form of private income was loudly condemned as usury. In modern times the pendulum has swung in the other direction. Interest is generally conceded to be earned income and its receipt by private persons is approved, while rent is often considered to be an unearned income. It is contended, as

a matter of fact, that there is little difference between the two types of income and that from the point of view of private individuals the one is about as likely to be earned as the other. And so the argument goes on and on.

QUESTIONS AND PROBLEMS

1. Distinguish between implicit and explicit rent; between economic rent and contract rent.

2. "Rent is a payment for the use of land and not for the using up of land." Explain.

3. Distinguish between the use price of land (or rent) and the purchase price of land.

4. If the general interest rate declined from 4 per cent to 2 per cent, how would you expect land values to be affected? Explain.

5. "The determination of rent for a single grade of land occurs under conditions known as 'composite demand.'" Explain and illustrate.

6. Explain rent determination for a single grade of land in the long run.

7. "Some of the conclusions which are usually reached in connection with rent determination for a single grade of land would almost certainly not be valid in the short run." Show whether you agree.

8. "The assumptions of the theory of distribution are likely to be realized in practice in connection with the determination of rents." Do you agree? Explain.

9. Explain the determination of rent as a differential surplus between two grades of land.

10. Is it possible for all grades of land used in a given field of production to command economic rents? Explain.

11. How will rents be affected by increases in the quantity of labor and capital available for use in conjunction with land? Explain.

12. "The rents received by landowners are unearned income from the social point of view and should be eliminated." Show whether you agree.

13. "If it were not for the operation of the Law of Diminishing Productivity, economic rent for land would not exist." Do you agree? Explain.

14. One furniture store in a certain city is located outside the high-rent district while another is located in the heart of the business section. Would you expect the difference in the rents paid by these stores to be reflected in the prices of their furniture? Explain.

15. "Landowners or tenants on high-grade land can afford to pay higher wages for labor than those on low-grade land." Do you agree? Explain.

16. If the government would no longer permit private individuals and companies to receive rent, how, if at all, would you expect the total quantity of land used in production to be affected? Explain.

17. "If the government imposed a tax of 50 per cent on economic rents, the burden of the tax would be felt by tenants and consumers rather than by landowners." Show whether you agree.

18. "If a city adopted a building code which restricted the height of all

buildings to 100 feet, rents would be seriously affected." Show whether you agree.

19. Suppose superhighways for automobiles were constructed for 50 miles in all directions from a certain city. Show whether you would expect rents within the city to be affected, and, if so, how.

See References for Further Reading at the end of Chapter XXII.

XIX

Interest

After having dealt with rent as the payment which is made for the use of land, it would seem logical to define interest as the payment which is made for the use of capital (capital goods), but there are several reasons why such a definition would be undesirable. In the first place, it is clear that capital goods as such are very seldom borrowed and lent, and interest payments do not often take the form of physical goods. It is ordinarily money, funds, purchasing power, or savings that we borrow and lend, and interest is usually a money payment. In the second place, when funds are borrowed and interest is paid, the borrowed funds are by no means always used to acquire capital goods. Since borrowed funds are sometimes spent for consumers' goods, or are invested in land, or may even be used to pay wages and other running expenses of business enterprises, it does not seem wise to establish too close a connection in theory between the payment of interest and the use of capital goods in production.

Finally, in those cases in which funds or savings are actually invested in capital goods of various types, it is difficult to conceive of the returns derived from the use of these capital goods by their owners, or the payments made for their use by other people, as interest. The enterpriser may borrow savings and agree to pay interest at a certain rate, but the return which he will realize from the investment of these funds in fixed capital goods may be either much larger or much smaller than the amount of interest which he agreed to pay for the use of the saved funds. The owner of fixed capital goods of superior quality has, in the short run, an advantage over the owner of fixed capital goods of inferior quality which is very similar to the advantage enjoyed by the owner of superior land as compared with the owner of inferior land.

THE NATURE OF INTEREST

In the long run, it is not difficult to reconcile the earnings of fixed capital goods with the rate of interest. The position of the owner of superior capital goods is very much like that of the owner of superior land

in the short run, but it is quite different in the long run. In the ordinary situation the owner of superior land will keep his advantage over the owner of inferior land even in the long run, for it is impossible to produce more land of superior quality. The owner of superior fixed capital goods will lose his advantage over the owner of inferior capital goods in the long run, because it is possible to produce additional capital goods of superior quality. As long as the rate of return which is expected from the use of such superior fixed capital goods is above common interest on the cost of producing these goods, there will be an incentive, under competitive conditions, to produce these goods in ever-increasing quantities. The process will be stopped only when there are so many of these capital goods in use that the rate of return to be expected from their use is no greater than common interest on the funds invested in them and is also no greater than the rate of return which can be expected from savings which are put into other fixed capital uses. Thus the rate of return derived from fixed capital goods of superior quality, which may be well above the interest rate in the short run, is transformed into common interest on the cost of producing such capital goods in the long run.

Interest Defined. As the above discussion has implied, savings or funds are not to be thought of as being tied up in particular types of capital goods in the long run. Savings in the long run are fluid or mobile and are capable of going into any field of production or type of investment. They are always in the process of distributing themselves among the various fields of production and types of investments in such a way that they will command the same rate of remuneration in each of their various alternative uses. This rate of remuneration is called interest. In other words, *interest is the payment which is made (per year) for the use of savings in the long run.* However, in defining interest in this way, we must bear in mind that business enterprisers often want to borrow savings in order to obtain capital goods or other agents of production, and that the ability of these enterprisers to pay interest and repay the principal sums borrowed will be closely related to the productivity of the capital goods or other agents of production which are acquired with the borrowed funds.

Net Interest and Gross Interest. An objection is frequently raised at this point. How, it is asked, can the economist talk about interest as a single rate of remuneration when it is obvious that many different rates of interest prevail at any given time? Thus, at any one time, businessmen may obtain short-term loans from commercial banks at 6 or 7 per cent interest, necessitous borrowers may secure funds from personal finance companies, pawnbrokers, or "loan sharks" at 3 per cent a month (36 per cent a year) or even more, corporations may borrow by issuing bonds at 5 or 6 per cent interest, the federal government may sell its bonds at 2 or 2½ per cent interest, insurance companies may pay dividends to policyholders on

the basis of an interest rate of 3 per cent, savings banks may pay 1 or 2 per cent interest to depositors, Brazilian or Peruvian bonds may bear 8 per cent interest, and call loans to brokers may be made by bankers at a rate of interest as low as 1 per cent or as high as 30 per cent.

The solution of this problem is relatively simple. These various rates of interest are what the economist calls "gross interest," while the rate in which the economist is chiefly interested is called "net interest." We mean by net interest the payment which is necessary to overcome the reluctance of savers to put aside funds out of current income and which represents the prospective usefulness of the funds or savings to the productive or consumptive borrowers. Gross interest is net interest plus a payment for the risk which is borne by the owners of the savings plus a payment for any administrative costs which are incurred in handling the savings and making and collecting the loans. If the owner of savings could be perfectly sure of receiving his interest and principal as promised, he would be content with an amount of interest which just compensated him for postponing the consumption of the amount of money income which he saved, but the different uses to which his savings may be put involve actually a varying degree of risk. The dishonesty of borrowers, the failure of businesses, fires, floods, breakdowns of machinery, the repudiation of debts by governments, and other things may prevent the owner of savings from ever receiving his interest or even his principal. Whether or not the individual risks are insurable, the owner of savings is not likely to entrust them to borrowers unless he can obtain a gross rate of interest which will both pay him for the use of the funds which he gives up and make it worth while for him to bear the risks involved in the uses to which the savings are to be put. Since the extent of the risk borne by the owner of savings varies significantly from one type of investment to another, we must expect gross interest rates to vary in similar fashion as a result.

A varying amount of administrative expense may also be involved in borrowing and lending activities. Savings are often entrusted to banks, which must maintain a considerable volume of productive facilities and a fair-sized personnel in connection with their lending activities. In general, the expenses involved must be covered by the interest paid by borrowers. In other cases the amount of administrative expense may be comparatively slight, but there is usually some such expense, and it tends to be covered by the gross interest rates which prevail in various types of loans. Since gross interest rates include payments for risk and administrative expenses, they depart from the net interest rate by an amount which varies significantly from one use of savings to another.

The Uniform Rate of Net Interest. The economist does not conclude that these differences in gross interest tend to be eliminated in the long run, for this result could not occur unless risk and administrative cost could

be eliminated or brought to uniformity in the various uses of savings. However, while gross rates of interest will vary from one use of savings to another even in the long run, economists hold that, underneath these different gross rates of interest, there tends to exist a single uniform rate of net interest under competitive conditions in the long run. This conclusion is based on two main points. In the first place, savings are all of one grade or quality, and we do not encounter the problem of different grades, as in the case of land or labor even in the long run, or as in the case of physical capital goods in the short run. Second, while savings are more or less immobile in the short run, they are highly mobile in the long run. Savings are inanimate and, unlike workers, they have no friends, home ties, children in school, or homes partly paid for, which might keep them fixed in one place or in one occupation. If savings will command more net interest in some uses of funds than in others or in some places than in others, after allowing for risk and costs of administration, it is to be expected in general that savings will move *in the long run* from occupations or places where net interest is low to occupations or places where net interest is high. The result is a tendency toward a single rate of net interest in the long run within a given market for savings.

Some people think that such a uniform net rate of interest tends to prevail for the United States as a whole. In fact, it has even been suggested that a uniform rate of net interest tends to exist between countries. Although questions of patriotism and difficulties of supervising the uses which are made of savings enter into the international picture, it may be that many people will put their savings into Peruvian bonds instead of into those of the United States government if they conclude that the high gross rate of interest on the former bonds, after allowing for the degree of risk involved, will furnish more net interest than the prevailing gross rate on the latter bonds will provide. Furthermore, the rather large direct investments which our individuals and companies have made in foreign lands and productive properties give some indication of the tendency of savings to pass from country to country in search of high net rates of return.

In this country, incidentally, the prevailing rate of interest on bonds of the United States government is considered a fair approximation of the current rate of net interest, since risk and administrative costs are supposed to be at a minimum in this use of savings. Where the interest rate in another use, as for example the interest rate on call loans, falls below the prevailing rate on government bonds, it is often supposed that the banks making these loans are sacrificing even a part of obtainable net interest in order to secure extreme liquidity and to be able to get their funds back on very short notice. We shall have occasion to refer again to the relations

between gross and net interest in presenting a diagram for interest determination in the long run.

INTEREST DETERMINATION IN THE LONG RUN

The Demand for Savings in Consumption. A fully developed demand and supply analysis may be presented in connection with the determination of the interest rate in the long run. The demand for savings is really a composite demand, like the demand for a single grade of land described in the preceding chapter, because savings are demanded by business enterprisers for use in production and by consumers for the purpose of obtaining economic goods for consumption. Strictly speaking, the demand for savings for productive use could be subdivided to show the demand for savings in each line of production, but this refinement would involve unnecessary complications and we shall deal with only two partial demands for savings —for use in consumption and for use in production.

How can consumers afford to pay interest for savings which are to be used in acquiring consumers' goods or services? Can consumers by borrowing acquire economic goods which they could not obtain without borrowing? The answer to the latter question is that they ordinarily cannot unless they are trying to cheat the seller of the goods or the lender of the savings. If a person buys an automobile on the installment plan and is able to save enough out of his income month by month to meet the "easy" payments and pay for the car completely within a year's time, it seems clear that, within the same length of time, he could even more easily save enough out of his income to buy the car outright at the end of the period. We say "even more easily" because the total amount of savings necessary will be less if the car is bought for cash than if it is bought on the installment plan. Some consumers may object to this analysis on the ground that they will save if it is necessary to do so in order to meet installment payments, but will not save in order to acquire a fund with which to make a purchase in the future. This objection may be accepted in cases in which individuals are so weak willed as to be unable to save without a finance company hanging to their coattails, but in other cases consumers do not often acquire goods by borrowing which they could not obtain without borrowing.

Since the borrowers who are under discussion at present are getting savings for purposes of consumption, the use of the savings will not furnish a money income from which the interest may be paid. It is therefore necessary to look to the utility of the economic goods acquired with the savings for the reason which accounts for the borrowing. At this point students sometimes jump to the false conclusion that consumers will borrow funds for consumption purposes as long as the utility of the goods to be acquired exceeds the interest which must be paid for the borrowed funds. If this

were true, the volume of consumer-borrowing would be extremely great, for it would be a peculiar economic good which did not have enough utility to cover a payment of 10 or 20 per cent of its purchase price. What the consumer really gets by borrowing is the use of an economic good now instead of later, and it is the *excess* of present over future utility which must be great enough to cover the necessary interest payment.

This excess of present over future utility is very real, for most persons prefer goods now to goods of the same quantity and quality at a later time. The rate at which a good now is preferred to the same good later on is called the rate of time preference, and it is on the relation between this rate and the rate of interest that consumer borrowing depends to a great extent. That is, if a consumer regards a good now as 10 per cent more desirable than the same good a year from now, and the necessary funds for purchasing the good now can be secured at an interest charge of 5 per cent for the year, he is likely to borrow for consumption. However, if the present good is considered only 3 per cent more desirable than the future good, and the funds required to purchase the good now cannot be secured at an interest rate of less than 8 per cent for the year, the consumer is not likely to borrow for consumption.

The rate of time preference varies not only from one individual to another at a given time and from one time to another for the same individual, but also from one economic good to another for the same individual at the same time. For example, if the individual goes to a furniture store tc look at radios and the salesman shows him a desirable specimen, saying, "Here is a nice radio which you can have for $100 in cash or for $200 in 10 easy payments," our customer is likely to parody Shakespeare in his reply, "Go to; I will wait until I can buy for cash, and in the meantime will rely on the loud playing of my neighbor's radio." However, if the consumer needs a major operation, cannot get free medical attention, must have $150 to pay for the operation, and is forced to go to a "loan shark" who will offer him $150 now if he will pay back $300 later on, the consumer may accept the proposition. He will reflect that he may not be physically able to work until he has saved the $150 and that, if he waits to try to save this sum, he may not even be around to take advantage of the surgeon's services.

In spite of all variations in time preference, certain conclusions may be reached in regard to the demand for savings to be used in consumption. If the rate of interest is high, borrowers for consumption are likely to take only relatively small quantities of savings because there will not be many economic goods for which the rate of time preference will be great enough to warrant the payment of such an interest rate. If the rate of interest is low, consumers are likely to borrow relatively large quantities of savings because there will be many economic goods for which the rate of time pref-

erence will be great enough to justify the payment of such a low interest rate. Thus the quantity of savings which consumers will borrow per year tends to vary inversely with the interest rate, in accordance with the Law of Demand, under the assumed conditions of competition (which would here include, of course, reasonable knowledge and information on the part of borrowers as to the size of the financing or interest charges which they actually have to pay).

The Demand for Savings in Production. Short-term borrowings by business enterprisers for productive purposes may also be explained on the basis of a simple time analysis. When the business enterpriser has goods produced but not yet sold, or goods sold but not yet paid for in cash, he is assured of having purchasing power after a relatively short period of time, such as 30, 60, or 90 days. In the meantime, however, he would be short of funds for paying wages, buying materials and supplies, and paying public utility rates if he did not borrow. By borrowing at a commercial bank on his own promissory note or on some piece of commercial paper furnished him by his customer in lieu of cash, he is able to obtain funds now instead of 30, 60, or 90 days from now. These funds will permit him to operate his business continuously and efficiently instead of by means of occasional bursts of activity. The advantage gained in this way is sufficient to make the enterpriser willing to pay interest for the use of present funds. However, the total quantity of short-term funds which business enterprisers would borrow would be expected to vary inversely with the interest rate. That is, larger quantities of funds would be borrowed at 5 per cent interest than at 15 or 20 per cent.

When the business enterpriser borrows savings on a long-term basis for the purpose of investment in fixed capital goods, the situation is somewhat different. The time element is present in such loans, of course, but the enterpriser is not seeking a simple time advantage. The issue is one of increasing the capacity and productivity of his enterprise and not simply one of having goods now rather than goods later. When the consumer borrows to buy a radio, the commodity purchased with the borrowed funds will not add to the borrower's money income, and he must pay off the loan plus interest out of the money income which he would have had anyhow. However, when the enterpriser borrows in order to construct a new wing on his factory and furnish it with necessary machinery and equipment, the commodities acquired with the borrowed funds will add to the money income of the enterprise during the life of the loan. If he waited to acquire the new capital goods for cash out of the income of his original enterprise, he might never get them. Consequently, in considering the advisability of such borrowing, the enterpriser must consider the addition which the new capital goods will make to the value-product of his enterprise in relation to

the cost of obtaining the necessary funds, and not simply the desirability of having goods now instead of goods later.

If the enterpriser is using very small amounts of capital in conjunction with given quantities of the other agents of production, he will probably be willing to pay interest at a rather high rate in order to obtain a few additional units of savings, because he will expect that the capital goods secured with these savings will add greatly to the productivity of his enterprise. If capital goods are already plentiful in his enterprise in relation to the given quantities of the other agents of production, he will be willing to pay only a rather low rate of interest in order to get additional units of savings, for he will expect the capital goods obtained with these savings to add comparatively little to the productivity of his business. In any given situation, he will tend to stop acquiring additional units of savings at the point after which the capital goods secured with another unit of savings would be expected to add less to the productivity of his enterprise than the amount of the interest payment necessary to secure the additional unit of savings.

Since each enterpriser would be assumed to function in the same general way, we can predict the nature of the entire demand for savings to be invested in fixed capital goods. If the rate of interest is high, business enterprisers will borrow only relatively small quantities of savings per year because it is only by keeping capital goods relatively scarce in their enterprises that they can maintain the marginal productivity of these goods at a level high enough to warrant the payment of this high interest rate. If the rate of interest is low, business enterprisers will borrow relatively large amounts of savings per year. The marginal productivity of the capital goods acquired with these large quantities of savings will not be high, but it does not need to be high in order to cover such a low rate of interest. Thus the quantity of savings which business enterprisers will borrow per year for investment in fixed capital goods tends to vary inversely with the interest rate, in accordance with the Law of Demand, under the assumed conditions of distribution.[1] Moreover, since the partial demands for savings behave in accordance with the Law of Demand, it is clear that the total demand for savings, which is merely a summation of the partial demands, must do the same. That is, the total quantity of savings borrowed per year for all purposes tends to vary inversely with the rate of interest.

The Supply of Savings in the Long Run. Having seen why borrowers for production or consumption are *able* or *willing* to pay interest for the use of savings, we must now ask why it is *necessary* that these borrowers pay interest. The answer is that, while a certain amount of saving would

[1] It should be remembered that one of these assumed conditions is that of stability of the general price level and purchasing power of money. In a boom period of business, when the price level is rising, businessmen may borrow larger and larger sums at higher and higher interest rates through time, but this has nothing to do with the situation which we are describing.

probably occur even if savers did not receive interest, the volume of saving necessary to produce equilibrium conditions in the market for savings would not be forthcoming in the long run if interest were not paid to the savers. Many, though not all, units of savings involve a rather definite cost to the savers—a cost which can be measured fairly well by a money price or interest rate, and a cost which savers would not bear unless such a rate of remuneration were obtainable.

Motives for Saving. According to this analysis, the desire to receive a reward in the form of interest is a very important motive for saving, but this is not to deny that there are many other motives for saving or that some saving would occur if interest were not paid. Even without interest many individuals would be inclined to lay aside certain amounts out of their money incomes to provide for the various financial emergencies which may come to anyone, to secure a life of luxury and leisure in future years, to provide for their heirs and dependents, or to obtain the power which the accumulation of large means may bring. Very rich people are said to save almost automatically because of the difficulty, if not the impossibility, of finding sufficient consumptive uses for their large incomes.

Various governmental units often collect taxes from the citizens and invest part of the revenue in fixed capital goods. Thus, in effect, they save for the individual citizens, although the latter never consciously decide what amounts are to be saved or what benefits are to be derived. Other units of savings presumably require interest but do not result from a voluntary choice on the part of the individuals whose funds are being saved. Such are the savings which are present when corporations reinvest the earnings of their businesses instead of paying dividends to the stockholders. Finally, it is sometimes contended that there are persons whose savings will tend to vary inversely rather than directly with the interest rate in the long run. If a man desired a future income of $5,000 a year from interest, he would have to save $200,000 if the interest rate were 2½ per cent but only $100,000 if the interest rate were 5 per cent.

Time Preference and Saving. These considerations show only that some saving would occur if no interest could be obtained by the savers; that some units of savings may not involve cost or sacrifice on the part of the savers, or even conscious decisions to save on the part of the individuals whose income is, in effect, saved; and that the savings of some persons might conceivably vary inversely with the interest rate, instead of directly. They do not show that the present total annual volume of savings would be made available to borrowers in the long run without interest or without sacrifice on the part of the savers, or that this volume of savings tends to vary inversely with the interest rate.

Except for such amounts of income as would be saved more or less automatically or without interest, the desirability to the individual of spend-

ing his money income in the period in which it is received, instead of at some future time, is considerable. Following the "bird-in-hand" philosophy, most persons value economic goods at present more highly than economic goods of the same quantity and quality at some future time. Enjoyments which have actually been experienced cannot be taken away later, but many things can happen to prevent the enjoyment of parts of one's income whose consumption has been postponed until a later time. Indeed, a person may have left this "sorry scheme of things" altogether and have passed on to a better world without getting an opportunity to consume income which he has saved. If the individual can look forward to a probable increase in annual income in the future, the desirability of spending his present income in the present is likely to be enhanced. Finally, there is the important fact that the money incomes of most individuals are so small that they are under considerable pressure to spend all their incomes in the period in which they are received in order to maintain a fairly satisfactory standard of living or even to continue to live.

While most individuals value present goods more highly than future goods of the same quantity and quality, saving requires that present goods be given up and future goods be accepted in their place. Except for the automatic and no-interest savings, if we approach an individual who needs or desires to spend his money income at present and ask him to lend us $100 in return for our assurance that we will return this sum to him intact after a couple of years, we may not ordinarily expect that our proposition will be accepted. We should be asking him in effect to give up something now ($100) in return for something of less present value ($100 receivable after two years). If we promise to return $110 or $120 after two years, in return for $100 now, the individual may seriously consider and eventually accept our proposition. The extra $10 or $20 is the premium or interest necessary to offset the individual's time preference, or his preference for goods now as compared with goods later.

Not only do many units of savings involve a time preference on the part of savers which must be overcome by the payment of interest, but it is also true that savings are subject to a principle or tendency of increasing time preference. That is, the greater is the sum which we ask an individual to save out of an annual or other income of given size, the higher is the rate of interest which we must pay in order to offset his preference for present as compared with future consumption. If a person had a money income of $500 per month, he might save $50 per month almost automatically or without interest. If we asked him to save a second $50 per month, it might be necessary to pay him 1 or 2 per cent interest. If we asked him for a third monthly installment of $50 of savings, an interest rate of 4 or 5 per cent a year might be required, and so on. In general, the greater the sum which an individual has already saved out of a given money income,

the higher must be the interest rate if we are to induce him to save any more in the same period of time. Thus in the total supply schedule for savings in the economic system in the long run, we must expect small total volumes of savings to be associated with low rates of interest per year and large total volumes of savings to be associated with high rates of interest per year. In a diagram the long-run supply curve for savings will start low on the left and move upward to the right.

The Illustration of Interest Determination. In Figure 34 we present an illustration of the determination of the net interest rate *in the long run* under the various assumptions of the theory of distribution. Both the demand curves and the supply curve in this diagram are drawn in terms of net interest rates. The actual gross interest rates which borrowers pay on the basis of time preference or productivity must include payments for risk and administrative costs as well as net interest, but we are assuming that varying amounts representing payments for risk and administrative costs have been deducted from these gross interest rates so that our demand curves are expressed in terms of net interest rates. In similar fashion, the interest rates in terms of which the supply curve for savings is drawn are net interest rates. They include only the payments necessary to overcome the time preference of the savers and disregard any payments which savers receive on account of the risk element or which anyone (not necessarily the saver) receives on account of administrative costs.

The demand curve d shows the demand for savings by borrowers who wish to use the savings productively. According to our previous analysis it must show that these borrowers would take relatively small quantities of savings per year at high interest rates and relatively large quantities at low interest rates. At each interest rate shown on the OY axis the amount of savings which would be taken by borrowers for consumption is added to the amount which would be taken by borrowers for production, and the result is the total demand curve D, which shows the demand for savings at all net interest rates by both classes of borrowers put together. Clearly the D curve on the diagram must draw away from the d curve at the lower interest rates since the borrowers for consumption will also take larger quantities of savings per year at low interest rates than at high interest rates.

The supply curve S indicates that small amounts of savings per year would be saved and offered for use at low interest rates. The larger the annual volume of savings required, however, the higher must be the rate of interest paid in order to overcome the time preference of the individuals who do the saving. Thus the supply curve is roughly similar to the long-run supply curve for a competitive industry operating under conditions of increasing cost. We must also note that the supply curve for savings starts at a point below zero with respect to the OY axis. This indicates that the

desire to provide for financial emergencies, and for other things previously mentioned, is so strong that most individuals would set aside certain minimum amounts of income even if no interest were paid, or even at negative interest if no better alternative were available. While it is not contended that many persons would accept negative interest in situations in which positive interest could be obtained, it is quite possible that individuals would pay interest, if necessary, for the safekeeping of their savings in order to have certain sums available when needed in the future. There

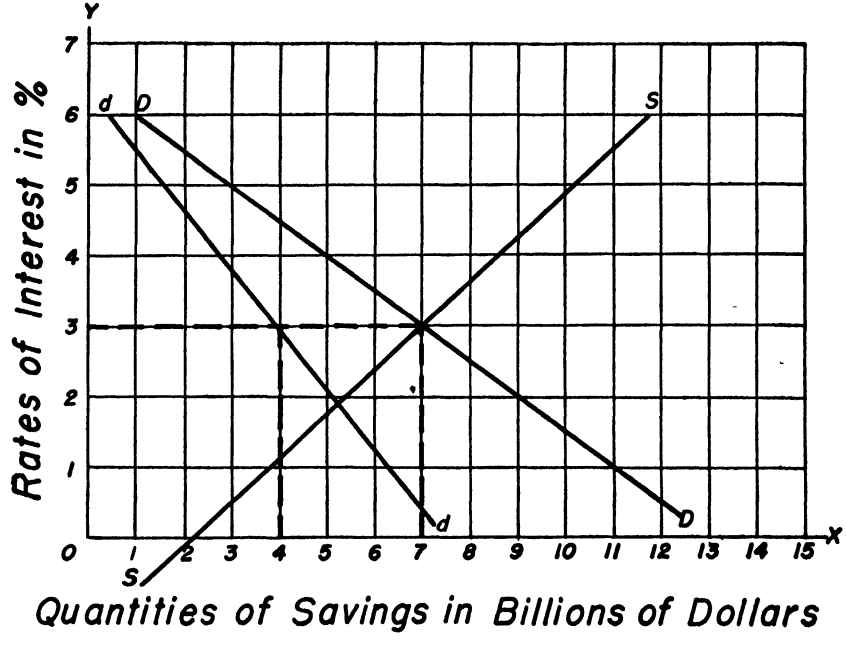

FIGURE 34.—Determination of the Net Interest Rate in the Long Run under Competitive Conditions

is a possibility, at least in the short run, that individuals would keep these savings in safe deposit boxes or domestic hiding places in order to have them always ready for use when needed, instead of offering them for investment, but our supply curve for savings is drawn on the long-run assumption that all savings which are made actually become available in the market for investment.

Under the demand and supply conditions shown in Figure 34 the rate of net interest would be 3 per cent a year in the long run under the conditions assumed by the theory of distribution. The sum of $7,000,000,000 would be supplied and demanded annually and, of this sum, $4,000,000,000 would go into productive uses while $3,000,000,000 would be used for pur-

poses of consumption.[2] It is not difficult to see why, under the assumed conditions, the rate of net interest could not be permanently higher or lower than 3 per cent. At the rate of 4 per cent, for example, borrowers would take only $5,000,000,000 annually, while savers would furnish between $8,000,000,000 and $9,000,000,000 annually, and competition of savers to find employment for their savings would tend to drive the rate of net interest down toward the level of 3 per cent. At a rate of 2 per cent, savers would offer only a little less than $5,500,000,000 per year while borrowers would be willing to take $9,000,000,000 annually, and competition of the borrowers to obtain the available savings would tend to force the net interest rate up toward the level of 3 per cent.

The Validity of the Assumptions of Distribution Theory. The discussion of the determination of interest in the long run has been conducted on the basis of the usual set of assumptions for the theory of distribution. If these assumptions are well realized in practice, the description of the determination of interest will afford a good approximation of the way in which interest is actually determined. However, since we must not make the fairly common mistake of taking assumptions to be facts, we shall now consider whether the various assumed conditions are likely to prevail in the actual market for savings. It is fairly likely that borrowers for productive purposes have a reasonable knowledge of market conditions in connection with savings, of the prospective marginal productivity of the agents of production to be acquired with borrowed funds, and of the rates of gross interest which enterprisers are paying in the various uses of savings. Borrowers for consumption are not usually nearly so well informed as are productive borrowers. They may be so ill informed concerning their various opportunities for borrowing that they borrow funds at high interest from personal finance companies which they could obtain from banks or other agencies at much lower rates. They are especially likely to be misinformed as to the exact rate of interest which they actually pay in borrowing for consumption. Individual savers are not usually aware of the alternative opportunities for using their funds, the degree of risk involved in various uses of savings, and so on, but the savers quite often do not handle the investment of their savings directly. Instead, the investment function is transferred to banks, insurance companies, and other institutions which are likely to be very well informed in these matters.

The actual number of borrowers and savers in the market should be large enough in general to satisfy the competitive assumption. That is, any saver can safely assume that he can save more or less than usual without affecting the rate of net interest which he can obtain, and likewise virtually any borrower can expand or contract his borrowing operations without

[2] All these figures are used merely for illustrative purposes, and it must not be presumed that they are accurate for the United States at any given time.

causing a change in the prevailing rate of net interest. Again, it is ordinarily true that, in practice, there is active competition among the demanders of savings to obtain the quantities of savings available in the long run, as well as among savers (or their agents) to take advantage of the available opportunities for the use of the funds. Organized attempts on the part of borrowers to restrict borrowing and keep the interest rate low, or on the part of savers to limit saving and keep the interest rate high, are quite uncommon. However, the bankers who operate as middlemen between the borrowers and savers have sometimes been accused of such organized activities.

The assumption that savings are mobile and capable of being transferred from one occupation to another or from one place to another is likely, as we have seen, to be very well realized in practice in the long run, though it is admittedly not worth much at any particular time. Substantial equality of bargaining power may not exist between borrowers and savers in actual practice, but it may exist in some cases between borrowers and the institutions which have charge of the investment or lending of savings. The assumption that the government does not interfere with the process by means of which the interest rate is determined, leaving this rate entirely at the mercy of demand and supply forces in the market, has been quite true in the more or less distant past. In recent years, however, governmental agencies of one kind or another have been interfering directly or indirectly with the gross interest rates which prevail in most, if not all, of the various uses to which savings are put in the United States.

Finally, we have the assumption that the borrowers of savings for production sell their finished products under competitive conditions. It is only under these conditions that the individual producer would be able to sell an additional amount of product without affecting the price per unit which he would receive for his original output, and hence could afford to hire additional units of savings as long as the interest payment did not exceed the value of the marginal product contributed by the agents of production to be acquired with the additional units of savings. Under conditions of monopoly or monopolistic competition in the market for finished products, the price per unit which can be obtained by an enterpriser for an economic good depends upon the quantity of the good which he produces and offers for sale. The hiring of additional agents of production and the sale of additional units of product involve the taking of a lower price per unit for the finished goods, both for the additional units and for those of the original output. Therefore an enterpriser under these conditions could afford to borrow additional units of savings and acquire additional productive agents only as long as the value of the marginal product of the agents exceeded the interest payment on the borrowed funds by at least enough to cover the loss of revenue which would result from the sale of the original output at a lower price per unit than formerly.

Since competitive conditions are seldom fully present in the markets for finished products, we need to observe briefly how our analysis would be affected by this situation. If some of the productive borrowers sell their products under competitive conditions and others under monopoly or monopolistic competition, the demand curve for savings on the part of productive borrowers as a whole will move downward more rapidly or sharply from left to right than our *d* curve in Figure 34. That is, a given lowering of the interest rate will cause a smaller increase in the quantity of savings demanded for production under these conditions than under those of pure competition in the sale of finished products. The total demand curve, which results from adding to the productive demand the demand for savings to be used in consumption, will be altered in the same way and the supply curve presumably will not be affected. Thus the equilibrium point will indicate a lower rate of net interest and a smaller annual volume of saving and borrowing under these conditions than in a situation in which the finished products of productive borrowers are sold under competitive conditions.

Since the various assumptions of the theory of distribution are not all realized completely or even to the same extent in the actual market for savings in the long run, it is difficult to evaluate exactly the theory of interest determination which has been presented. However, on the whole, this theory probably furnishes a fair approximation of the actual process of the determination of interest in the long run.

INTEREST, SAVINGS, AND INVESTMENT IN SHORT PERIODS OF TIME

Our discussion of the demand for and supply of savings, and of the determination of interest, has been stated in terms of the long-run period. From this point of view and because of what we know about the expansion of human desires for finished commodities and services, it does not seem probable that the available quantity of capital goods, or of savings for investment in capital goods, will ever become excessive in relation to the uses which we should like to make of these capital goods and funds. If this is true, then the long-run analysis of supply should deal, as it has, with the obstacles to saving and capital formation and with the necessity of paying interest in order that adequate quantities of savings may be made available. Moreover, it is necessary to conclude that the amount of savings supplied on the market will vary directly with the interest rate in the long run.

However, in relatively short periods of time, and especially amid the cyclical fluctuations which occur in economic activity, the situation with regard to interest, savings, and investment may be quite different. Some economists contend, for example, that consumption tends to increase, but not as fast as the output of goods and the money incomes of individuals, in a period of business prosperity and expanding production. In such a situation, in the absence of important new inventions or overconfidence on the

part of businessmen with regard to the future, the prospect of securing adequate earnings from further investments in new capital assets becomes progressively less favorable. As the prospective marginal productivity of capital investments declines, the inducement to invest declines also and may readily reach zero. However, the rate of interest at which savings or capital funds may be obtained never reaches zero, although it may decline, and investment activity may virtually cease.

In the long run we should expect a decline in the rate of interest to result in retarded saving and increased consumption, but nothing of the sort may happen in a relatively short period of time. The reason is that the willingness of individuals to expend their money incomes in consumption—or their propensity to consume, as it is sometimes called—is quite stable in the short run. If this is true, we cannot expect the market supply of savings, or funds withheld from consumption, to vary directly with the interest rate in such a period, and saving may go on at a considerable rate in spite of a decline in the rate of interest and the disappearance of the demand for savings to be invested in new capital assets. This situation, as we shall see later, has important implications for the emergence and continuation of depressed conditions of business. For the present we are content to note that, in relatively short periods of time, there may be a very considerable problem of "oversaving."

This analysis may be entirely valid, but we should not allow it to convince us that the real problem in connection with the supply of savings in the long run is to find some use for the savings which people are determined to make rather than to induce adequate saving. In the short run the individual income receivers of our economy may have a rather stable propensity to consume, but in the long run the propensity to consume may keep pace very well with the expanding output of goods. Just as the fact that a monopolist may operate at a loss in the short run does not deny him the possibility of operating at a profit in the long run, so any failure of the quantity of savings to vary directly with the interest rate in the short run does not mean that the quantity of savings will not vary directly with the interest rate in the long run. Individuals may save or withhold from consumption considerable amounts of money income in the short run only to see these sums lie idle in the form of cash balances, but there would be no point in such behavior in the long run. In that period, funds which are saved tend also to be invested, and there is not likely to be any problem of oversaving.

QUESTIONS AND PROBLEMS

1. "Interest is a payment for the use of capital goods." Discuss.

2. "The return derived from the use of capital goods may be greater than the interest rate in the short run, but it can scarcely remain greater in the long run." Explain.

3. "The income status of the owner of superior capital goods is quite similar to that of the owner of superior land." Show whether you agree.

4. Distinguish between net interest and gross interest.

5. "It is silly for the economist to maintain that a single rate of interest prevails in the United States when everyone knows that many different rates of interest actually exist in this country?" Do you agree? Explain.

6. "By borrowing funds consumers can obtain economic goods which they could not obtain without borrowing." Discuss.

7. What is the consideration which consumers should bear in mind in deciding whether it is worth while to borrow funds for purposes of consumption? Explain.

8. "Short-term borrowing by business enterprisers for productive purposes, like consumer borrowing, can be explained on the basis of a simple time analysis." Show whether you agree.

9. "In borrowing savings on a long-term basis for the purpose of investment in fixed capital goods, the enterpriser is not seeking a simple time advantage." Do you agree? Explain.

10. "The demand curve for savings to be used productively has the same general characteristics as the demand curve for funds to be used in consumption." Explain.

11. The United States Steel Corporation wants to borrow $100,000,000 by means of a bond issue. Show why it might be able to pay 5 per cent interest annually for the funds, and why it might find it had to pay this rate.

12. "The motives for saving are many and various." Explain.

13. "The theory of time preference is merely a formal statement of the bird-in-hand philosophy." Discuss.

14. "A loan without interest is fundamentally similar to a sale at a price below cost." Do you agree? Explain.

15. Show how a long-run supply curve for savings is derived from the theory of time preference.

16. Draw and explain the diagram for the determination of the interest rate in the long run.

17. "Since any prevailing rate of interest gives some savers more inducement than they need in order to save and permits some borrowers to acquire funds at a rate lower than they would be willing to pay, it is ridiculous to refer to such a prevailing interest rate as an equilibrium rate." Show whether you agree.

18. Are the conditions assumed by the theory of distribution likely to prevail in the actual market for savings? Explain.

19. How is the theory of interest affected if many productive borrowers sell their products under conditions of monopoly or monopolistic competition? Explain.

20. "In relatively short periods of time the situation with regard to interest, savings, and investment may be quite different from that which exists in the long run." Explain.

21. Discuss the problem of "oversaving" in short and long periods of time.

See References for Further Reading at the end of Chapter XXII.

XX

Wages

Labor has been defined as human energy expended for the purpose of acquiring income, and *wages are the payments which are made (per year) for the use of labor in production.* For the purposes of our discussion of wages we include in labor all human energy which is expended in production without regard to the question of whether the energy is expended in mental or physical work or in some combination of the two. In discussing wages we encounter once more the complicating factors which troubled us in connection with the determination of rent. That is, the total quantity or number of workers is divided into different grades or qualities, and each grade is ordinarily capable of being used in several different lines of production. Moreover, it may be possible to use different grades of labor for a given type of work. After considering first the theory of the determination of wages in connection with the services of workers of a single grade or quality, we shall pass on to a study of the differences in wages which prevail between the various grades of workers.

WAGE DETERMINATION FOR A SINGLE GRADE OF LABOR

The Productivity Analysis. Later on we shall deal with the problems which are involved in describing the characteristics of the various groups or grades of labor. For the present we shall assume that it is possible to distinguish a certain grade of labor, such as unskilled labor, and shall turn to the question of the unit which should be used in discussing the determination of wages for labor of a given grade. It is customary to use the individual worker as the unit of labor and to assume that a given grade of labor is composed of numerous workers of the same efficiency. Actually, however, differences in efficiency do exist from one worker to another within a given grade of labor, and the theory of wages is weakened to some extent by dealing with wages on a per-worker basis. Therefore we shall state the theory of wages in terms of units of labor service, or *number of hours of work of given intensity by workers of a given grade.* This makes it possible to recognize that one unskilled worker may furnish many more

(or fewer) units of labor service in a year's time than another worker of the same grade.

Because the discussion of the general theory of distribution in Chapter XVII was conducted largely in terms of labor we are already familiar with the general nature of the demand for labor of a given grade, but the demand analysis should be reviewed here. If an enterpriser desires to use unskilled labor in his productive establishment and secures a number of units of labor service of this grade which is very small in the relation to the fixed quantities of the other agents which he employs in production, he will find that both the total product of his enterprise and the product per unit of labor service will be relatively small. Much of the power generated in the plant will be wasted, much of the machinery and other facilities will be idle most of the time, and each worker will have to perform a number of different tasks or functions.

As more units of unskilled labor service are used along with the fixed productive agents, the quantity of labor will become relatively more sufficient, better use can be made of the fixed productive facilities, and the workers can concentrate on individual specialized tasks. The result will be both an increasing total product and an increasing average product per unit of labor service in this stage of operations, for each additional unit of labor service will add to the total product an amount of product greater than that produced on the average by previous units of labor service (or, in other words, the marginal product of labor will be greater than the average product). The enterpriser will scarcely wish to stop acquiring units of labor service so long as additional units will increase both the total product of the enterprise and the product per unit of labor service.

Sooner or later, however, the tendency to diminishing productivity per unit of the variable productive agent will assert itself. The specialization of labor by tasks can be carried only so far and the fixed productive facilities can be operated more and more effectively by an increasing number of units of labor service only up to a certain point. Once the point of diminishing productivity has been passed, the total product of the enterprise will go on increasing but the average product per unit of labor service will decline as additional units of labor service are employed. In other words, the marginal product of labor will be declining and will be less than the average product of labor. It is somewhere in this stage of increasing total product but declining average and marginal product of labor that the enterpriser will want to stop taking on units of labor service. In other words, the enterpriser will have no incentive to go beyond the point of absolutely diminishing returns, for beyond this point each unit of labor service can add nothing to the total product or will even cause it to decline.

The last unit of labor service which is employed is least important

and least productive when combined with its share of the other productive agents. This unit is therefore called the marginal unit of labor service, its product is called the marginal product, and the money value of this marginal product is the marginal productivity of units of labor service of this grade in this establishment. If the enterpriser sells his product under conditions of competition, this marginal productivity sets the limit on the wage which can be paid for the marginal unit of labor service and hence for each other unit of labor service of this grade, since the various units are alike and interchangeable in the productive establishment.

The Demand Analysis. In converting this discussion of diminishing and marginal productivity into a statement of demand conditions, we may say that, if the wage rate for this grade of labor is high, our enterpriser will use a relatively small number of units of labor service of this grade in his business in conjunction with given amounts of the other productive agents. It is only by keeping these units of labor service relatively scarce in his business that the enterpriser can maintain the marginal productivity of the labor at a level high enough to cover this high wage rate. If the wage rate for this grade of labor is low, the enterpriser will demand a large number of units of labor service for use in conjunction with the given amounts of the other productive agents in his establishment. If the labor is used in large quantities, its marginal productivity will not be high, but it does not need to be high in order to cover this low wage. At any wage rate which may prevail, the enterpriser will employ units of labor service in his business down to the point where the addition of another unit of labor service would cause the marginal productivity of the labor to fall below the wage rate which must be paid per unit of labor service. In other words, within the effective region in which the total product of the enterprise is rising but the product per unit of labor service of this grade is falling, the curve of marginal productivity for units of labor service of this grade may be taken as the individual competitive enterpriser's demand curve for this labor.

As we have just seen, the number of units of labor service of a given grade which the individual enterpriser will demand for use in his business varies inversely with the wage rate in a given market at a given time. Since other enterprisers may be expected to behave in the same fashion as the enterpriser with whom we have been dealing, we may conclude that the total demand for labor of this grade in all of its various alternative uses will also obey the Law of Demand and that the total demand curve to represent this demand in a diagram would start high on the left and move downward toward the right. Since the demand for labor of a given grade is really a composite demand, we should plot separately the demand for this labor in each of its various alternative uses and add these demands together horizontally to arrive at the total demand curve for the labor.

The diagram which would result would be similar, from the point of view of the demand situation, to that for a single grade of land, as shown in Chapter XVIII.

Short-Run Supply Analysis. In thinking of the number of units of labor service of a given grade which may be offered on the market for use in production, we may decide at once that, in any short period of time, there will be a fixed number of persons in existence in the population who are capable of supplying units of labor service of the given grade. However, this fact should not lead us, as it did some economists of many years ago, to the conclusion that there is a "fixed supply" of labor of a given grade in the short run, for the number of persons who actually offer themselves for employment may vary considerably even though the number of persons capable of furnishing units of labor service of a given grade is fixed. In general, we should expect more individuals to offer themselves for employment at high wages than at low wages. At any rate, during the period of World War II, the market supply of labor of various grades in the United States was increased considerably under the influence of high wages and other incentives.

The number of persons offering themselves for employment is not the only possible source of variation in the supply of labor of a given grade in the short run. That is, some at least of the workers who come out for employment may be willing to work harder (with greater intensity of effort) at a high wage rate than at a low wage rate, or may be willing to work longer hours at a high wage rate than at a low wage rate, and either of these factors would affect the number of units of labor service furnished. It may be true, as some economists have contended, that some workers would respond to high wages in a different way, would want to work less days or hours at a high wage rate than at a low one, or would even put forth less effort for high wages than for low ones. However, our general conclusion is that the number of units of labor service of a given grade which workers will offer on the market tends to vary directly with the wage rate in the short run, if we take into account the intensity of labor and the hours of labor as well as the number of workers seeking employment. On this basis the short-run supply of labor of a given grade would be represented in a diagram by a supply curve which started low on the left and moved upward toward the right.

Long-Run Supply Analysis: Subsistence. In the short run, and especially in a situation of national emergency, it may be possible to induce changes in the number of people who offer themselves for employment out of the total number capable of furnishing units of labor service of a given grade, in the hours per day or week which the workers will spend at their jobs, and in the intensity of effort put forth by the workers. In the long run, however, these factors may be considered to be relatively stable or at least

as factors which will change only slowly and gradually. Consequently, in discussing the supply of labor of a given grade in the long run, chief emphasis is placed on a factor which does not change at all in the short run; i.e., population, or number of persons in existence who are capable of furnishing units of labor service of a given grade. In order to predict long-run supply conditions, it would be necessary to establish a relationship between wage rates and numbers of workers of the grade in question. If we want the number of unskilled workers to increase in the long run, should we pay high wages to workers of this grade in the meantime, or would low wages be more effective? Unfortunately for the purposes of the theory of wages it does not seem possible to give any reliable answer to questions of this kind.

Some economists have tried to establish a positive relationship between wage rates and numbers of workers in the long run. Starting with the assumption that workers tend to receive a subsistence wage or a wage just high enough to enable them to live, remain able to work, and replace themselves in the laboring population, they reasoned that, if workers came to receive an unusually high rate of wages, the workers would raise larger families or at any rate more of their children would survive and eventually take their places in the laboring population, and the number of workers would increase. In similar fashion, if the wage rate remained exceptionally low for any great length of time, the workers would have smaller families or at any rate fewer of their children would survive long enough to become part of the labor supply, and the number of workers would eventually decrease. If this analysis could be depended upon, the long-run supply curve for units of labor service of a given grade would resemble the long-run supply curve for an industry operating under increasing cost. That is, the higher the wage rate in the meantime, the larger would be the number of workers of this grade in the long run, and the larger would be the number of units of labor service of this grade in the long run, and vice versa. However, almost no one today thinks that there is a direct relationship between wage rates and the number of workers of a given grade in the long run.

Long-Run Supply Analysis: Standard of Living. There are some writers who believe that some sort of inverse relationship exists between the wage rate and the number of workers of a given grade in the long run. It is said that, if we pay workers of a given grade unusually high wages for any considerable period of time, they become accustomed to a higher scale of living than that which they formerly enjoyed and they will struggle to maintain this high scale of living even if it means marrying later in life or restricting the size of their families. On the other hand, it is alleged that nothing depresses workers so much or is so likely to make them improvident, imprudent, and careless about the future as being compelled by

force of circumstances to accept wage rates which make it impossible for them to maintain the scale of living to which they have become accustomed. On this basis one might predict a long-run supply curve for units of labor service of a certain grade which would resemble the long-run supply curve for an industry operating under decreasing cost. That is, high wage rates would be associated with small numbers of workers and units of labor service in the long run, and low wage rates would be associated with large numbers of workers and units of labor service. However, there is very little evidence to support this point of view.

Long-Run Conclusions. The upshot of the whole matter seems to be that the relationship between wage rates and numbers of workers in the long run is not at all direct and close. The growth of numbers or population is the result of a large number of forces, many of which are noneconomic or only partly economic in character. Most parents do not rear children for profit or financial gain and, even if they wanted to do so, it is clear that the size of the prevailing wage rate would not furnish a very effective stimulus. The wage rate for a given type of labor may be high at present, but since it takes many years to bring children up to the age at which they may join the working population, the wage rate has a chance to change hundreds of times before new units may be added to the labor supply. Children who were brought into the world so that they might contribute wages to the family exchequer might found families of their own long before they could be of much assistance to their own parents. Moreover, a man who wished to have a large family of sons who could be, say, steelworkers might find himself the father of daughters instead.

Thus, for various reasons, there seems to be no way to express at all satisfactorily a relationship between the long-run supply of labor of a certain grade, depending as it does on changes in the numbers of workers, and wage rates. Without such a relationship there is no useful long-run supply analysis, and it is impossible to present a diagram illustrating the determination of the wages of labor of a given grade in the long run. About all that can be said is that, under competitive conditions, the wage rate for units of labor service of a given grade in the long run depends upon the marginal productivity of labor of that grade, and that this in turn depends upon the number of workers in the group or grade, if we take as given the intensity of labor, the hours of work, and the proportion of potential workers who offer themselves for employment. All the workers in a given group or grade of labor tend to receive the same rate of wages *per unit of labor service* in the long run, regardless of the particular occupations in which they expend their productive energies. However, the total amounts of money which they receive may vary considerably from one occupation to another on the basis of various attractive and unattractive conditions of work.

THE VALIDITY OF THE COMPETITIVE ASSUMPTIONS

Knowledge of the Market. Even the broad conclusions just stated depend upon the various assumptions of the theory of distribution in the long run and may be considered valid as descriptions of actual economic life only if these assumptions are realized or nearly realized in practice. This consideration requires that we now turn to a re-examination of these assumptions. In the first place, is it likely that workers and the employers of workers actually have the knowledge of various market conditions which is required by the assumption of competitive conditions? The employers of labor may have a fairly good general knowledge of the sources of supply of labor, the wages paid to workers of a certain grade in various occupations, the prospective marginal productivity of units of labor service in their own businesses, and other matters. The workers, on the other hand, are often woefully ignorant of the value of their services to their own employer, their other opportunities for employment, the requirements of other jobs or occupations, and the wages which they could obtain in other occupations. The extent of the workers' knowledge of the market probably improves as we pass from low-grade to high-grade groups of workers, but it is almost never so great as is assumed under the heading of competitive conditions.

Numbers of Workers and Employers. Another assumption of competitive conditions is that there are large numbers of both buyers and sellers in the market. If this is true in the market for labor, both the supply of labor of a certain grade to the individual employer and the demand for labor of a certain grade to the individual worker will be perfectly elastic. That is, the individual employer must be able to buy a larger or smaller number of units of labor service of a certain grade than usual without affecting the wage rate which he has to pay, and the individual worker must be able to offer and furnish a larger number of units of labor service than usual without affecting the wage rate which he receives. To the extent that employers and workers function as individuals, there can be little doubt that this condition is largely realized in practice in the labor market as a whole and in the markets for large groups or grades of labor. However, there are some high grades of labor which include only a few workers. Indeed some high-grade workers may be said to possess almost a complete monopoly of the specific varieties of labor service which they furnish. Some small grades of labor may be demanded by only a few employers, or possibly by a single employer. In all such cases the assumption under discussion does not hold.

Independence of Individual Workers and Employers. A third assumption of competitive conditions is that both employers and workers function as individuals so that employers compete actively to obtain workers and workers compete actively to obtain employment. This condition is relied

upon in theory as a force which will bring wage rates back to a marginal productivity level in the long run when they depart from this level in short periods of time. If some employers pay less than a marginal productivity wage to their workers, other employers are supposed to bid higher rates for these workers and compel the first employers to raise their wage rate or lose their workers. In similar fashion, if some workers of a certain grade are paid a wage rate above their marginal productivity, other workers of the same grade are supposed to offer to work for lower wages, and such competitive bidding is supposed to cause wages to fall back toward the marginal productivity level.

This assumption is often classified as one which is partly true in practice although affected by strong opposing tendencies. At times the competitive bidding of employers for workers or of workers for employment can be seen clearly in actual markets for labor. Among the opposing forces, as far as employers are concerned, we may list such things as formal or informal combinations of employers to lower wages or keep them from being raised, the custom of small enterprisers to follow the lead of large enterprisers in matters of wage and labor policy, ignorance or inertia on the part of employers, fear of reprisals, and so on. On the workers' side, the chief opposing force is found in the organization of workers into unions through which the workers compete for employment and wages as a group and not as individuals, though ignorance and inertia also play a part.

If some of the workers in a given group or grade of labor form a union and place severe restrictions upon membership, they may be able to obtain a wage rate higher than that which used to prevail on the basis of marginal productivity for the group as a whole. Enterprisers who cannot avoid the use of this union labor will go on hiring it, but they will react to the high wage rate by limiting the number of workers they employ, so that the marginal productivity of the union workers becomes great enough to cover the wage rate. With employment restricted in these enterprises, more workers of this grade than formerly must be employed in other enterprises and industries so that their marginal productivity and wage rate become lower than formerly. While the wage rates of union and nonunion workers may both equal the marginal productivity of the respective workers, the outcome of the determination of wages is sharply different from that which would have existed in the absence of organization on the part of the workers.

The situation is somewhat different if all, or practically all, of the workers of a given grade or group of labor are organized. Of course, if the workers were formerly paid wages which were below the level of marginal productivity, the labor organization may be useful in bringing about a rise in wages to the marginal productivity level. However, if the labor organization is able to set a wage rate for a whole grade or group of labor

which is above the marginal productivity level for labor of that grade, employers of the labor will respond by hiring a smaller number of the workers than formerly. In such a case the high wage rate will come to be covered by the marginal productivity of the smaller number of workers employed (in conjunction with given quantities of the other productive agents), but the employed workers will gain only at the expense of their fellows who are unemployed.

Thus the general conclusion is that, if workers of a certain grade are already being paid on the basis of marginal productivity, a labor union can obtain still higher wages for some of the workers only by driving others into unemployment. However, this conclusion depends on the assumption that the marginal productivity of the labor is influenced only by the proportions in which the labor is combined with other agents of production, and cannot be affected by the wage rate itself. In some cases, if the workers receive higher wages than formerly, their individual productivity may increase, either because they become better able to work than formerly or because they become willing to do more work than formerly. If the productiveness of the workers increases under high wages so that they come to earn what they are paid, employers may not find it necessary to reduce the number of workers which they employ and no unemployment need result. Again, the necessity of paying unusually high wages to the workers may spur some employers to improve the organization and operation of their enterprises so that the workers can earn the high wages which they demand. In this case, too, the conclusion that some workers can obtain wages higher than their former marginal productivity only by driving other workers into unemployment may not follow.

If a group of employers in an industry conspire to reduce the wage rate paid to workers of a certain grade, workers who are tied to the industry and cannot move to other occupations may continue to work even though the new wage rate in their employment is lower than that which used to prevail on the basis of marginal productivity for the grade of labor as a whole. Since the conspiring employers will be paying the workers a wage rate which is lower than the workers' marginal productivity, they may realize large financial gains for the time being. In the long run, however, the employers may find it difficult to retain their workers at the wage rate which they have decided to pay. That is, the workers who are immobile in the short run may become able to move to other occupations and industries in the long run or they may counter the employers' organization with an organization of their own.

The Mobility of Labor. The theory of the determination of wages also depends on the assumption that workers are mobile as to places and occupations, and will move from places or occupations where wages are low to other places or occupations where they are high. This assumed

condition, like most of the others, is only partly realized in practice. In short periods of time, the immobility of workers as to places is notorious. Among the factors which keep workers from moving are attachments to particular localities, the desire to live near friends and relatives, the cost of moving families and establishing them in new locations, the losses which might have to be taken in selling houses or other investments in the old localities, the effect of constant moving on the education of children, and ignorance of opportunities in other places. However, high-grade workers are likely to be more mobile between places than low-grade workers, and all grades of workers are likely to be more mobile over long periods of time than in short periods. The assumption of mobility between different occupations in a given locality is better realized in practice than that of mobility between places. The chief factors preventing the movement of workers in a given locality among occupations which are open to the members of a given labor group are ignorance and inertia. Of course, workers cannot move from occupations open to a certain grade of labor to occupations which are carried on by members of a higher labor group because, as we shall see later, the various groups of labor are largely noncompeting.

Equality of Bargaining Power. Another assumption of the theories of distribution and wages is that individual buyers and sellers are approximately equal in bargaining power. This is usually not true in practice between individual workers and individual employers. The individual worker may need a particular job very badly while the individual employer seldom cares very much whether or not he employs a given worker, especially if the worker is in one of the lower groups or grades. The individual worker often lacks knowledge about labor opportunities even if he could go elsewhere to seek them, and he usually does not have the financial reserves which would enable him to hold out for high wages. Lack of equality of bargaining power could not be of great importance in a market in which workers had complete knowledge of market conditions, were sure of being employed at something or other, were completely mobile, and were subject to a highly competitive demand on the part of employers, but in actual markets it may be a factor of considerable importance. When workers combine in unions, their bargaining power is increased, but it is not necessarily equal to that of the individual employers. Moreover, if employers also combine, the effectiveness of workers' combinations may well be partly destroyed.

Lack of Governmental Interference. The theory of the determination of wages assumes that wages are freely determined by demand and supply conditions and by the activities of workers and employers, and that governmental units do not interfere with the process. This assumption is largely untrue in so far as it concerns actual markets for labor in the United States in recent times. The existence of relief or work relief for workers

who do not have private employment may certainly be expected to have an effect upon the competition of workers for private employment and upon their willingness to accept private employment at low rates of wages. The federal government's encouragement of union organization on the part of the workers and its championship of the "rights of labor" tend to affect the relative bargaining power of workers and employers and the activeness of competition of workers for jobs. Legislation discouraging labor by women and children will affect the marginal productivity of adult male workers and the wages which they receive. In wartime, ceilings on wages, the freezing of workers in essential jobs, and attempts to drive workers out of unessential work may interfere still further with the ordinary process of wage determination.

The setting by the federal government or state governments of minimum wages and maximum hours in certain ranges of occupations may also have an effect on the determination of wages. Minimum wages in some cases may merely bring the wages of workers of low bargaining power up to the level of marginal productivity, but in other cases these wages, as set by law, may exceed the marginal productivity of the workers affected by them. If the minimum wages set at a level above marginal productivity apply only to some of the occupations which are open to the members of a labor group or grade, employment tends to be restricted in the regulated occupations, while more workers than formerly find their way into the unregulated occupations and wages tend to fall there. On the other hand, if these high minimum wages apply to all the occupations which are open to the members of a labor group or grade, employers in general are likely to use this labor sparingly, in combination with given quantities of the other productive agents, in order to keep its marginal productivity high enough to cover the minimum wage. The result will be that some workers of this grade will be able to receive the high minimum wage while other workers of the same grade will be unemployed.

These conclusions assume that the marginal productivity of the workers is a function of the proportions in which labor is combined with other factors of production and cannot be affected by the wage rate itself. In actual practice, some workers may show an increase in productiveness as they receive high rather than low wages, either because they become better able to work or because they become more willing to work. Moreover, some employers, under the pressure of having to pay high minimum wages, may improve the organization and operation of their enterprises to such an extent that the workers can show an increase in productivity. If the workers, for any reason, increase their productivity under the high minimum wage so that they come to earn what they are paid, then the minimum wage has no necessary adverse effect upon employment.

The Influence of Monopoly and Monopolistic Competition. Finally, the theory of wages depends upon the assumption that the employers of labor sell the products of their businesses under competitive conditions. Only under these conditions is the enterpriser assured that the price of his product is independent of the amount of product which he individually puts on the market so that he can hire units of labor service as long as the marginal physical product of labor times the price of the product is adequate to cover the wage rate. If the finished products are sold under conditions of monopoly or monopolistic competition, the price which the enterpriser can obtain for his product will vary inversely with the amount which he offers for sale in a given period of time. He can therefore afford to hire an additional number of units of labor service only if the value of the prospective additional product, when discounted for the decline in price which must be taken on the originally contemplated output, is adequate to cover the wage rate at which the additional labor can be hired. The existence of large elements of monopoly or monopolistic competition in the markets for finished products will lower, other things being equal, the wage rate at which all of a given number of units of labor service of a certain grade will be employed.

Appraisal of the Theory of Wages. After examining all of these factors, what shall we finally say of the general theory of wages? Since the assumed conditions which surround the theory of wage determination are not completely present in practical situations, it is clear that the theory cannot be considered an exact explanation of the process by which wages are determined in actual markets for labor. On the other hand, almost no theorist today contends that this theory of wages gives anything more than a rough approximation of a long-run tendency. The fact that many extraneous factors often keep the theory from working perfectly in practice does not mean that the tendency described is not present at all in actual markets. Moreover, even if we were sure that some factors would always prevent the working out of wages equal to marginal productivity in practice, it might still be desirable, before plunging into the thousand and one complications which exist in reality, to begin the study of wages by examining the manner in which wages would be determined under certain simplified conditions or assumptions. It does not seem likely that more can be learned about wages by concentrating attention solely on the exceptions and complicating factors than by proceeding from the statement of a general theory to the introduction of other matters. The chief danger in developing a general theory of wages is that the theorist, having made certain assumptions for the purpose of developing the theory, will later forget that his assumptions were really assumptions and come to believe that they are conditions which actually exist in the markets for labor and that his theory

is therefore perfectly attuned to reality. This danger we have earnestly tried to avoid.

GROSS WAGES AND NET WAGES

Differences in Cost of Training. Under the assumptions of the theory of wages, the rate of wages received by the workers of a given group or grade tends to equal the marginal productivity of the units of labor service which these workers furnish in the long run. Moreover, the units of labor service furnished by workers of a given group or grade tend to command the same *net rate* of wages in the long run in all the various occupations which are filled by members of the group. However, the *gross wages* paid per week or per year may differ to some extent from one occupation to another within the same labor group even in the long run and under the assumptions of the theory of wages. Net wages include merely the wage rates which are paid because of the marginal productivity of the units of labor service furnished by the members of a labor group. Gross wages include net wages plus the payments which are made to workers on account of the risk, disagreeableness, cost of training, or other features of certain occupations.

For example, suppose the members of the same labor group could become either carpenters or electricians, but that the former occupation required almost no training of a formal sort while the latter occupation required a long and expensive course of training. If gross wages were the same in the two occupations, then under the assumptions of the theory of wages the workers would become carpenters and not electricians. Carpenters would become so numerous that their gross wages would fall and electricians would become so scarce that their gross wages would rise. These changes would stop at the point where the gross wages of electricians exceeded those of carpenters to such an extent that, after allowing for the greater cost of training electricians, net wages based on marginal productivity were the same in both occupations. It is not to be supposed that such a precise adjustment is actually reached in practice, but it would be attained under all the assumptions of the theory of wages. Even in practice some adjustment of gross wages on the basis of differences in the cost of training or other costs of entering certain occupations is likely to be made.

Risk and Disagreeableness. Under the assumptions of the theory of wages, similar variations in gross wages tend to exist on the basis of the varying degree of risk or disagreeableness in different occupations. Because of the greater risk to health or life itself which prevails in these occupations, workers with pick and shovel in mines, or in caissons under high air pressure in the construction of bridges or of tunnels under water, tend to receive higher gross wages than pick-and-shovel workers who function on the surface of the ground. Such high gross wages are somewhat like

the royalties which are paid in connection with the use of mineral lands. That is, they are payments for using up or depreciating, instead of merely using, the agent of production in question. High gross wages are also likely to exist, under the usual assumptions of the theory of wages, on the basis of the disagreeableness of certain occupations, quite apart from any risk or danger involved. The saver of funds may not care at all whether his funds are used in operating commercial flower gardens or in running glue factories, but the worker, who always has to deliver his units of labor service in person, may have a clear-cut choice between these two opportunities for the employment of his labor if the gross wages are the same in both cases.

Psychic Income. Other occupations are supposed to be so pleasant and chock-full of psychic income that, unless the workers in them sacrifice a part of the wages which they could obtain in alternative occupations, all members of the labor group will flock to the occupations with the many psychic advantages. The occupation of college teaching is supposed, popularly, to be one of this sort. Let us suppose that young persons originating in the same labor group could be brought up to be either doctors or college professors, but that average gross wages are lower for college professors than for doctors. What does the college professor get from his occupation which may compensate for his lower wages? For one thing he gets short hours of duty. He teaches eight to twelve hours per week, while the doctor ordinarily must be on duty for a greater number of hours per week and in times of epidemics or other emergencies may work almost twenty-four hours per day for short periods, sleeping between calls and snatching a bite to eat when he can get it. Of course the college professor is expected to do a certain amount of work besides meeting his classes, but his nonteaching schedule will probably be fairly flexible and he may be quite certain that no epidemic of, say, economics will come along.

The college professor also enjoys regular hours of duty. He can be sure that no one is going to call him out at three o'clock in the morning to deliver a lecture on the theory of wages or on any other subject, while the doctor may be called out at almost any hour of the day or night. Both occupations are dignified and hold a high position in the community, but the college professor has all his patients come to him and thus works in pleasant surroundings, enjoys association with other persons who are much like himself, and has the privilege of training the best young minds. He continues to enjoy the college atmosphere and can, if he wishes, remain a college boy at heart all his life. He has a long summer vacation (too often without pay), while the doctor may hesitate to snatch even a couple of weeks away from his practice. The professor is given a certain income and is expected to accept the students who may come his way, while the doctor must attract his patients as best he can and depend for his income

on the fees which he can charge. Thus the case is usually presented, but members of the college teaching profession often suggest that the psychic advantages of their occupation are not so great, nor the money income so small, as is commonly pictured.

Once again it may be noted that we are not suggesting that, in practice, gross wages are always accurately adjusted to compensate for risk, disagreeableness, or psychic income, but such adjustments would be made perfectly under the conditions assumed in the theory of wages, and some adjustments are made in actual economic life. However, it sometimes happens in practice that the only workers who can be found to take risky or disagreeable jobs are so weak in bargaining power or so deficient in other respects that in some cases gross wages may actually be lower in these jobs than in others which are safer and more agreeable.

Other Factors. Many other factors are sometimes alleged to affect gross wages. Some jobs are blind alleys with no hope for future improvement in the status of the worker, while in others the worker may anticipate advancement later on. In connection with his position the ambassador to a foreign country has to meet many expenses of a sort which an ordinary business executive would not have to worry about. Some workers must buy their own uniforms and food, and rent a place to live; others, in addition to money wages, are furnished food, clothes, and living quarters by their employers. Some workers must buy goods from their employers at high prices in company stores or live in company houses at high cost while others are free to buy goods from their employers at a discount. Some workers have no chance to make supplementary earnings, while others can earn additional income in various ways. Under the assumptions of the theory of wages in all of these pairs of cases the workers named first would receive higher gross wages than the workers named second. Sometimes it is added that gross wages tend to be higher in seasonal occupations, such as coal mining or the building trades, than in other occupations, in which employment is relatively steady throughout the year. This is undoubtedly true as far as gross wages per day or per week are concerned, but it should not be true of the annual wages in which we are most interested.

The differences in gross wages among the occupations which are open to the members of a given group of workers are sometimes called "equalizing differences," since, under the assumptions of the theory of wages, they would leave the net wage status of all the workers in the group the same. If the workers in the different occupations got the same gross money and real wages, the net advantages of the various occupations would be unequal rather than equal, and workers would shift from some occupations to others. The stable condition would be that in which gross wages were unequal, but net advantages and net wages equal among the various occu-

pations within a certain labor group. However, the differences in gross wages in practice do not measure accurately the advantages and disadvantages of the various occupations nor do they leave all the workers in a group in precisely the same position with respect to net wages. These results would be obtained only if the assumptions of the theory of wages were actually realized in the markets for labor.

DIFFERENCES IN WAGES BETWEEN GROUPS OF WORKERS

The Noncompeting Groups of Workers. We have just been dealing with differences in *gross wages* among the various occupations which are filled by members of a given group of workers; now we turn to differences in *net wages* which prevail *between* the various groups or grades of workers. As was suggested earlier in this chapter, the total available number of workers in our economic system does not consist of a single large mass of homogeneous units, but is divided into groups or grades of workers which are usually called noncompeting groups. The number of these groups and their exact identity are difficult but not necessary to establish. Important for our purposes are the fact that there are such groups and the fact that they are noncompeting. The groups are called noncompeting because it is difficult, if not impossible, for the members of one group to qualify for occupations which are open to members of a higher group. This difficulty of upward movement through the groups often extends to the children and grandchildren of the members of the lower groups. The reasons which account for it will be examined later.

Let us consider first the groups of labor as they are commonly presented: (1) The group at the bottom of the heap is usually called unskilled labor. It is composed of workers who offer physical strength in production rather than intelligence or training. Most industries require some workers of this kind. They are found in manufacturing, in farming, in the building industry, in public service industries, and in many other places. Almost anyone in good health can do the work. The group is large and the average level of wages is low. Members of the group often go to work at an early age, reach their maximum wages early in their years of work, and suffer declining earnings later on.

(2) Workers in the next group are often called semiskilled, since they are neither skilled nor unskilled. Their work does not require great physical strength, and may need some intelligence or a small amount of training. Sometimes mentioned as members of this group are streetcar motormen, bus drivers, chauffeurs, some miners, some factory workers whose work in tending machines is quite routine and monotonous but requires some intelligence and responsibility, and routine office workers such as filing clerks. This list is intended to be merely suggestive and is by no means complete.

(3) In the middle of the classification is the group which is composed of skilled workers. This group includes carpenters, plumbers, bricklayers, electricians, garage mechanics, machinists, glass blowers (formerly), many railway employees, and a host of other types of workers. Their work in general requires a fair amount of intelligence and quite a little training, but not usually great physical strength. The group represents the aristocracy of the laboring class and in the past was responsible for most trade-union organizations.

(4) Above these three groups all is confusion. Sometimes a fourth group is set up to comprise business and professional workers of ordinary ability. Here we would find many managerial and office employees, doctors, lawyers, schoolteachers, salesmen, advertising men, small storekeepers, and a welter of other persons who cannot be classed in the lower groups and who are not in the ranks of the extremely well to do. As usually set up in theory, this group is something of a catch-all affair.

(5) At the top of the heap is usually found a group called the "upper crust," or "captains of industry," or something similar. It is composed of the tycoons of industry, the moguls of finance, the eminently successful doctors, lawyers, and other professional workers, the brightest stars of Hollywood, the home-run kings and other stars of professional sport, and others. It is another rather heterogeneous group.

Criticisms of the Theoretical Groups. We cannot support strongly either the number of groups in this classification or their make-up. Each of the bottom three groups may be made up of large numbers of almost homogeneous units of labor, but the same thing cannot be said of the two upper groups. These upper groups are not really groups at all but combinations of many small groups. In some cases there are very few workers in these small groups and sometimes individual workers enjoy something which closely approaches a personal monopoly over a certain kind of labor service. The actual number of labor groups is therefore not five but a number very much larger. Moreover, if the groups are to be really non-competing, they should be set up in such a way that the movement of workers from one occupation to another within each group is relatively easy, while movement between the groups is very difficult. This also is not always true of the groups as they are usually described. Both the fourth and fifth groups are composed of business and professional workers, and all that a doctor, lawyer, motion-picture actor, business manager or other worker in the fourth group has to do, in order to move up into the fifth group, is to make more money. This is quite different from what a ditch-digger would have to do in order to become a lawyer. Moreover a small storekeeper might find it as difficult to become a college professor as would a bricklayer, although the storekeeper and teacher are usually classified

in the same group, while the teacher and bricklayer are placed in different groups.

Why the Groups of Workers Are Noncompeting. In spite of the fact that the usual classification of groups of workers is far from satisfactory, it will not be necessary to go further in investigating the subject, since the important thing in connection with the theory of wages is the existence and noncompeting character of labor groups rather than their exact number and composition. When we say that groups of workers are noncompeting we mean that they are ordinarily noncompeting in an upward direction and not in a downward direction. That is, a plumber could probably do unskilled work if there happened to be any financial incentive for him to make the change, but the unskilled worker would experience great difficulty in becoming a plumber, and still more in becoming a college professor, however great might be the differences in wages between these occupations.

The reasons why an unskilled worker could not become a doctor are fairly obvious. He is likely to lack the ability required in the medical profession and he is not likely to have any opportunity to acquire the necessary training because he probably has to work as much as he can in order to make a living. But why does the economist argue that the unskilled worker's son would also encounter great difficulties in trying to become a doctor? Let us consider the situation for a moment. The first question is whether the son will have enough native ability and intelligence to qualify for the medical profession. It is quite possible that he will, but it is not the most probable thing in the world. If he has the ability, what then? He must go to school, and we must wonder whether his home influences will be such as will encourage him to be a good student, whether he will have adequate food and clothes, and glasses or aids to hearing if he needs them, and so on. After completing grade school, he must go to high school, and this may require better clothing and the purchase of books. Will he be able to stay in school past the age at which he can successfully evade the truant officer? Possibly so, but the pressure is great on the children of the poor to get out of school at the age of sixteen, or whatever the legal age is, take a job at "good" wages, and help support his family. The boys who do this often marry early in life, found their own families, and lose forever their opportunity for advancement.

After high school, our young man must go through three years of premedical work in college and then through four years of the study of medicine. In view of the difficulty of these studies, imagine his chances of completing them on his own power without financial assistance. If he succeeds, he will have to serve a year or so as an intern at very little salary at some hospital or other before he is ready to hang out his shingle as a doctor. And after going into practice for himself, he may have a long wait for patients. His chance of developing a profitable practice may be

reduced because he lacks the family and social connections which assist some young doctors. Now we are not contending that these difficulties are insuperable, but they are great and serious. Individuals do rise from the ranks to positions of affluence and power. It is possible to begin one's career as a newsboy and eventually become president of the United States Steel Corporation, if not of the United States, but these events make the headlines when they happen. They would not be front-page news if they were very common. And opportunities to rise from the ranks do not become greater as the economic system of a country grows older.

Wage Differences between the Labor Groups. On such considerations we rest our conclusion that there are labor groups which are largely noncompeting in character. We now turn to a discussion of wage differences between the groups. The average annual rate of wages undoubtedly increases as we ascend through the groups. This does not mean that every person in one group gets a higher wage than every person in a lower group. Some electricians may get greater wages than some lawyers and some routine officer workers may make less than some unskilled workers, but the average wage is higher in each group than in each lower group. The differences in wages from top to bottom between the groups are very great. Some captains of industry receive wages of $500,000 or $800,000 per year, while other full-time workers make as little as $500 per year. A worker who receives an annual wage of $500 would, of course, have to work a thousand years to make as much as the $500,000-a-year man receives in one year. Moreover, those individuals with the higher wage incomes are likely to have various property incomes in addition to wages. This is seldom true of the workers with low wage incomes.

Why do differences in wages prevail between the groups of labor? The reason is not that one large group is more important than another large group as a whole, because any one large group may be regarded as indispensable as a whole. We must find our answer in terms of marginal productivity, under the assumptions of the theory of wages. That is, if doctors receive higher wages than unskilled workers it is because the loss of productivity to society would be greater if one doctor were eliminated than it would be if one unskilled worker were eliminated. There is a difference in importance per unit of labor or labor service between the groups. These differences in importance or marginal productivity depend in turn upon differences in numbers. Many persons are qualified to function as unskilled workers while comparatively few persons are qualified to function as doctors. If the qualities of our population should change so that almost no one remained strong enough to do unskilled labor while almost everyone became intelligent enough to become a doctor, we might see unskilled workers riding around in their limousines while doctors would be obtainable at a dime a dozen.

The differences in numbers between the groups depend in turn on the noncompeting character of the groups. If upward movements through the groups were easy to accomplish, we should expect workers to move up from one group to another until net wages became equalized throughout the groups, but a combination of hereditary and environmental conditions largely prevents such movements. We shall make no attempt to decide whether the hereditary or environmental influences are more important. Some people believe that hereditary qualities are more important, and argue that ability, like murder, will out. According to this point of view, a person may be born in any given set of environmental conditions and will nevertheless rise to the top of the heap if he has sufficient native ability. On the other hand, some persons believe that we are currently wasting a large part of our managerial, professional, and artistic ability because environmental conditions chain individuals to the group or class of society into which they were born and never give them an opportunity to come to the top. They agree with the poet's sentiment to the effect that

> Full many a gem of purest ray serene
> The dark, unfathom'd caves of ocean bear;
> Full many a flower is born to blush unseen
> And waste its sweetness on the desert air.

The humble economist can suspect only that the truth lies somewhere between these opposing points of view.

Our discussion has shown that differences in wages between the noncompeting groups of labor rest first on differences in marginal productivity between the groups, second on differences in numbers of workers between the groups, and finally on the factors which render the groups noncompeting. The differences in net wages between the groups, under the assumptions of the theory of wages, would measure accurately the differences in marginal productivity between the units of labor service furnished by the various groups, and might be considered as "rents of ability" for the workers in the various groups. Under the actual conditions of labor markets, we know only that wage differences exist between the groups and cannot expect that these differences are always precisely adjusted on the basis of marginal productivity.

Other Wage Differences. There are many other differences in wages besides those which exist in gross wages among the different occupations filled by the members of any given group of labor and those which exist in net wages between the various groups of labor. However, we have examined the wage differences which are most important from the viewpoint of economic theory. In a country such as the United States there are

often differences in wages for given occupations from one region or part of the country to another. Sometimes these differences are apparent rather than real, because an occupation which goes by a given name may include somewhat different functions from one part of the country to another or because there are differences in price levels from one region to another so that there are no differences in real wages as great as those which prevail in money wages. Where the differences in wages from one part of the country to another in given occupations are real, they suggest that workers of certain kinds are scarcer in one region than another in relation to the other productive agents. In view of the well-known immobility of many types of workers between places, these differences in wages between regions may be quite permanent in character.

QUESTIONS AND PROBLEMS

1. "The problems which must be faced in explaining wage determination are essentially the same as those encountered in dealing with the determination of rent." Explain.

2. "The theory of wages is weakened to some extent by dealing with wages on a per-worker basis." Why?

3. Explain the dependence of the demand for labor of a certain grade on diminishing and marginal productivity.

4. "The number of units of labor service of a given grade which the individual enterpriser will demand for use in his business varies inversely with the wage rate in a given market at a given time." Explain.

5. "The number of units of labor service of a given grade which workers will offer on the market tends to vary directly with the wage rate in the short run." Show whether you agree.

6. "The continued payment of high wages for labor may be expected to lead to an increase in the number of workers in the long run." Discuss.

7. "The scale of living of labor of a certain grade is the result of past and present wages, and it may not be expected to influence the supply of that labor in the future." Do you agree? Explain.

8. "The relationship between wage rates and numbers of workers in the long run is not at all direct and close." Show whether you agree.

9. "The assumed conditions of distribution theory are less likely to be realized in practice in the market for labor than in the markets for the other agents of production." Do you agree? Explain.

10. "If the wages of workers equal marginal productivity, there is nothing that labor unions can do about raising wages." Show whether you agree.

11. "If the wages of workers already equal marginal productivity, minimum wage laws can raise wages for some workers only by throwing others out of employment." Do you agree? Explain.

12. Distinguish between gross and net wages.

13. Why might a test pilot receive higher gross wages than an ordinary commercial pilot for an airline? Explain.

14. "Differences in gross wages between occupations within a given group or grade of labor should be considered as 'equalizing differences.'" Explain and illustrate.

15. "Differences in wages between bricklayers and ditchdiggers are the result of other forces than those which may bring about differences in wages between bricklayers and electricians." Do you agree? Explain.

16. "There are exactly five noncompeting groups or grades of labor and the lines between them are well drawn." Discuss.

17. Why are certain groups of labor noncompeting? Explain.

18. How do you account for the differences in wages which prevail between the various noncompeting groups of labor? Explain.

19. "The noncompeting groups of labor are artificial rather than natural and would be eliminated if all individuals were given an opportunity to obtain education and training." Show whether you agree.

20. How would you account for the fact that "machinists" may receive $7 per day in Chicago and $5 per day in New Orleans? Explain.

See References for Further Reading at the end of Chapter XXII.

XXI

Profits

The discussion of profits is often a troublesome part of the theory of distribution. The student who has carefully followed the discussion of the four preceding chapters may well wonder why any examination of profits is necessary. Land, capital, and labor (including management) are useful in production, and through their cooperation the entire national income in each year is produced. For their productive services the owners of these agents of production receive rent, interest, and wages. These shares in distribution should exhaust the national income which was created by the cooperation of the productive agents. Why, then, should there be any additional income, known as profits, to consider? Who receives such income, and why?

THE NATURE OF PROFITS

Profits as All Residual Income. Much of the confusion which has existed in the discussion of profits has resulted from the different ways in which profits have been defined. Too often profits are thought of as all income remaining in the hands of the business enterpriser or firm at the end of the year after the money expenses of the business are paid, but it should be clear that such total residual income above money outgo is not a new and distinctive kind of income or share in distribution. Rather it is for the most part a welter of familiar types of income or distributive shares. The individual enterpriser has probably contributed land or capital or both to his business. If so, before deciding that he has any profits he should deduct from his total residual income a rate of remuneration for such agent or agents, equal to what he could have obtained for the agents if he had turned them over to some other enterpriser for use or to what he would have had to pay for the use of the agents if he had obtained them from someone else.

In similar fashion, since he has contributed his own abilities and services to the business, before counting his profits he should deduct a rate of remuneration for himself equal to what he could have obtained if he had hired himself out as a manager or to what he would have had to pay

for a manager of comparable ability. To be sure, the individual enterpriser may conclude that he is worth all the income which he has left over after paying money expenses, but whether or not this is good accounting, it is not good economics and it would cut the other way in years in which the enterpriser had little if anything left after paying his money expenses. At any rate, the individual enterpriser should figure up his profits only after deducting from his total money income all the implicit expenses or costs of operating his enterprise.

In the case of the corporation, since the managers are ordinarily hired outright, their wages will appear as money expenses of the business. However, they may receive a share in the profits of the enterprise and thus may be paid more than the competitive value of their productive services. From total residual income, after the payment of interest on bonds, before counting profits the corporation should deduct interest at the competitive rate on the funds which the stockholders have invested in the business, and the profits, if any, should be augmented by adding in any amounts paid to managers in excess of the competitive value of their productive services. It follows that most dividends to stockholders should be considered as interest. It is a case of self-delusion when a stockholder speaks of a $1 dividend as a "profit" on a share of stock in which he has invested $100, when he could obtain $5 per year from an investment of similar size in a corporate bond. Dividends on stocks should be considered profits only to the extent that they exceed interest at a competitive rate on the funds which the stockholders have invested in their businesses. With this fact in mind we realize that corporate profits are not nearly so common as they are popularly supposed to be.

The term "gross profits" is sometimes applied to the entire money income of an enterpriser or firm after money expenses are paid, but, as we have seen, gross profits are a composite return and few, if any, valuable conclusions can be reached in connection with such a return. Most of its parts are really costs of production, as the economist defines costs, and have already been accounted for in our study of rent, interest, and wages. However, it may be said that gross profits are absolutely necessary for any business which uses agents of production for whose use no money payment needs to be made to persons outside the business. Unless gross profits were made such an enterprise could not continue in business in the long run.

Profits as Net Residual Income. If we wish to examine profits as a separate type of income, we must mean by profits only *the excess money income received by the business enterprise or firm over and above the amounts necessary to remunerate all the agents of production, both those owned inside the business and those owned outside, at their full competitive rates.* Such profits are made only when the individual enterpriser or corporation has money income left over after paying all expenses and after

allowing for the remuneration of all agents of production which are furnished by the owner or owners of the business. Though common enough under monopoly or monopolistic competition, they are, as we shall see, relatively uncommon under competitive conditions in the markets for finished products.

Since profits as we have defined them are surplus income, they are not a cost of production and are not necessary to the successful conduct of a business in any period of time, short or long. If any enterprise could be sure that it would never receive less than cost of production for its products, it would never need to receive more. We see also that profits, as surplus income, are not earned by anyone. It is impossible to *earn* profits. Profits are a chance return and result from factors which are beyond the control of the individual business enterpriser or firm. If they are received by the business enterpriser, he gets them not for anything which he does in connection with production but because he is the residual agent of production and is the last to get paid off. He hires the other agents of production and contracts to pay them certain stipulated sums for their services in production. He must obtain a payment for his own productive agents and services out of whatever is left after paying the other agents of production. If, after the other agents are paid, the amount of money income which remains is more than enough to remunerate the enterpriser for his own agents and services, he retains the excess income. If the amount which remains is too small, he must get along as best he can with less than a fair return for his own agents and services.

The owners of any agent of production could be residual and receive profits when they occur. Landowners might hire labor, capital, and managers, furnish the necessary land, and rely on residual income for their rent. In such cases they would receive the profits if any were made. Capitalists could hire labor, land, and management, furnish capital, and depend on residual income for interest, plus profits, if any. This is substantially what happens today in the case of the corporation. Finally even workers might hire land, capital, and management, furnish labor service, and obtain their wages, plus possible profits, out of the residual income of the business. Profits are unnecessary, unearned, and unimputable income, and tend to be received by the owner of the last agent of production to be paid, which is in many cases the business enterpriser or furnisher of managerial labor service.

THE CAUSES OF PROFITS UNDER COMPETITION

In examining the causes of profits we shall find it helpful to remember the extent to which profits are likely to exist under competitive conditions in the markets for finished products. Clearly, in the long run under competitive conditions, no profits will exist for a whole industry. The price

of the economic good produced by such an industry tends to be equal to minimum average cost of production per unit in the long run and, while this cost of production includes "gross profits" or a return for the agents of production owned inside the various firms, it contains no element of profits as we have defined them. The firms in a competitive industry may make either positive or negative profits in short periods of time, but the industry tends to expand or contract in size and productive capacity in the long run under these conditions until no profits are made. Equally clearly, under our analysis, individual firms in a competitive industry will make no profits in the long run, for all the firms in such an industry will have the same minimum average cost per unit of product in this period. At any given time, some firms may have better agents of production than others, but in the long run all firms will come to have agents of production of the same superior quality if these agents of production are reproducible. If the superior agents of production are not reproducible, their value in the market or the value imputed to them by their owners will be bid up by the competitive process of capitalization in the market until there remains no net advantage to be derived from using the superior agents.

In the short run, individual firms in a competitive industry may make net profits, but they are not nearly so common as is usually supposed. If one firm receives more income than another because of using superior agents of production, the additional income is considered to be earned by these agents and is in the nature of a rent and not profits. Individual firms in a competitive industry make profits in the short run as a result of chance or risk factors beyond their control, and the factors which tend to cause profits for individual firms in a competitive industry will often cause such profits to exist for most of the firms in the industry in this period. Thus it is possible for a whole industry to receive profits in the short run under competitive conditions. Under conditions of monopoly or monopolistic competition in the markets for finished products, profits are common enough.

Profits as a Reward for the Efficiency of the Enterpriser. Many explanations of the existence of profits have been advanced. Before settling down to our own analysis of the causes of profits, we shall consider two explanations which, though popular, seem rather unsatisfactory. Some economists, after describing profits much as we have, conclude that one of the leading causes of profits is the individual ability or efficiency of the enterpriser. Such a conclusion not only is false but also does not follow from the description of profits. To be sure, the able, efficient enterpriser should and does make more money than the inefficient, incompetent one, and the difference in income may be called a rent of superior business ability. However, the difference between the enterprisers is found in "gross profits" and not in net profits.

The simple facts are these: Costs of production include full payments for the efficiency in production of all the productive agents, under competitive conditions, and business enterprisers or managers are included in these productive agents. If one individual enterpriser is more efficient than another, before figuring his profits he should allow himself greater wages of management, and if full allowance is made for his efficiency, he is likely to find that his profits are no greater than those of the less efficient enterpriser. In the case of the corporation, if one manager is more efficient than another, he should be paid a larger salary or wage of management by the corporation. If such wages of management measure accurately the relative efficiencies of the various managers, no corporation can make net profits and attribute them to the efficiency of its management. Thus in any business, if costs of production include full payments for the efficiency of enterprisers or managers, it is impossible to account for profits in terms of this same efficiency on the part of the enterprisers or managers.

Business experience also indicates that profits are no respecters of the abilities of business enterprisers. In some years net profits are general throughout a competitive industry. Almost anyone who can get into the business can make a profit. In other years practically every enterprise will show a net loss if it really takes all its costs of production into account. If an efficient enterpriser makes a large profit in one year and a considerable net loss a year or so later, how can we account for the change? Surely it is not logical to suppose that the enterpriser lost his efficiency over such a short period of time. He may have worked harder in one year than in the other, but if this is true, the chances are that his greatest efforts were put forth in the year in which he made the loss. Thus it is impossible to account for the presence or absence of profits through time in terms of the efficiency of the enterprisers.

However, in the case of the corporation, the whole relationship between wages of management and profits is often obscured. Some corporate officials may receive salaries which are actually smaller than those they could obtain from other companies, but they receive large additional amounts of income in the form of bonuses which vary from year to year with the success of the business. In other cases only straight salaries are paid, but these salaries are large enough to include a share of the profits in good years and a slice of the stockholders' interest in poor ones. The fact that profits, when made, may be shared with corporate officials suggests to some people that these officials have produced the profits, but to us it indicates the superior bargaining position of these officials rather than any responsibility for the existence of profits.

There is one case in which it might be possible to attribute profits to the ability or efficiency of the enterpriser, at least in theory. It is conceivable that an enterpriser may actually be more able and efficient when working

4. "The owners of business receive profits because of their residual position and not because of any function which they perform. Any agent of production could be residual and receive profits when they occur." Do you agree? Explain.

5. "The leading cause of profits is unusual efficiency on the part of the business enterpriser." Show whether you agree.

6. "The whole relationship between profits and wages of management is frequently obscured in practice." Explain.

7. "Profits may be accounted for satisfactorily as a reward for risk bearing on the part of the business enterpriser." Do you agree? Explain.

8. "There is a genuine connection between risks and profits." Show whether you agree.

9. "Profits result from the dynamic character of economic activity and would not exist if conditions were static." Discuss.

10. "Profits under competition result from the favorable outcome of risks which are beyond the control of the business enterpriser." Explain and illustrate.

11. "It is easy to account for the existence of profits under noncompetitive conditions." Explain.

12. Discuss "predatory activities" in relation to profits.

13. "The average rate of profits over the United States as a whole would probably be close to zero in the ordinary business year." Show whether you agree.

14. "If the government should impose a 100 per cent tax on profits, the result would be virtual confiscation of the values of stocks and bonds since corporations could no longer pay interest and dividends on these securities." Do you agree? Explain.

15. "It would be much better to keep profits from occurring than to tax them when they appear." Discuss.

16. "Governmental enterprises can sell goods at lower prices than competitive enterprises because it is not necessary that the former enterprises make profits." Show whether you agree.

See References for Further Reading at the end of Chapter XXII.

XXII

Personal and Noncapitalistic Distribution

According to the theory of the distribution of income which we have been discussing, the rate of remuneration received by the owners of a particular grade of a productive agent tends to be determined by demand and supply, in the long run and under competitive conditions, so that it equals the marginal productivity of the grade of the productive agent in question, or the marginal contribution which this productive agent is able to make to the exchange value of commodities and services. It is assumed, of course, that the units of a particular grade of a productive agent are so nearly alike that they may be interchanged in production and that they are able to contribute to the production of a number of different commodities and services with varying exchange values.

If only a few units of the particular grade of a productive agent are available, the rational allocation of productive resources requires that these units be reserved for their most important uses (most important, that is, from the point of view of creating exchange value and not necessarily from that of social utility or welfare). If a large number of units of the agent of production are available, they can be employed fully only if some of them are devoted to relatively unimportant uses—that is, to the production of economic goods which have comparatively little exchange value. Under any given supply conditions for the productive agent, it is the exchange value which is created by the agent in its least important use which is important in determining the rate of remuneration which will be received by the owners of all units of the agent. Since these units are alike and interchangeable, the loss of any unit whatsoever would deprive us, in the long run, only of the productivity of the agent in its least important use. That is, the uses of the agent would be reorganized so that the only use lost would be the one "off the end" or the least important or marginal use.

In consumption, if I consume a certain number of gallons of water in a day and I am deprived of those units which would ordinarily be used for drinking, I shall move the units which would normally be used for sprinkling the lawn up to the more important drinking use, and the only use of

water which I shall lose will be its least important use. This same analysis is applicable to the use of productive agents, and it is impossible to impute greater importance or productivity to any unit of a given grade of a productive agent than that of the marginal unit.

Long before any unit of the productive agent would be devoted to the fifth or sixth most important use, it would usually be profitable to devote more than one unit of the agent to the most important use. In fact, we should use a unit of the agent for even the second most important use only when it appears that the first unit devoted to this purpose will create more exchange value than would another unit devoted to the most important use, and so on down the line. The production of the various economic goods would be carried so far that the productivity of the last unit of the agent in the most important use would be equal to that of the last unit of the agent in any other use, and even in the last or marginal use of the agent. Thus we see that there is both an intensive and an extensive margin in the utilization of a given grade of a productive agent and that the marginal productivity of the agent will tend to be the same at both margins, in the long run and under competitive conditions.

These observations complicate the explanation to some extent, but they do not affect the essential conclusion, which is that all owners of units of a given grade of a productive agent tend to receive a rate of remuneration which depends upon the productivity of the agent in its marginal or least important uses, in the long run under competition. Bearing in mind possible reflex influences of rates of remuneration on the quantities of the productive agents which will be available in the long run, when we apply the same analysis to all other grades of the same productive agent and to all the grades of other productive agents we have a general theory of distribution which suggests how the money income (and indirectly the real income) of a capitalistic economy will be apportioned among the owners of the various grades of the productive agents.

The general theory of the distribution of income is based upon a considerable number of assumptions. Specifically, the theory assumes that (1) there are many demanders and suppliers of each productive agent; (2) the users of any agent compete actively for the available units of the agent, and the owners of the agent compete actively for the available opportunities for its employment, with no combinations, organizations, or conspiracies on either side of the market; (3) the owners of any productive agent and those who use it in production are reasonably well informed concerning market conditions which affect the agent; (4) each agent of production is mobile between occupations and, if possible, between places; (5) there is equality of bargaining power between the suppliers and demanders of each productive agent; (6) the government does not interfere in the distributive process and the rates of remuneration for owners of productive agents are determined

exclusively by demand and supply factors in the market; and (7) the enterprisers who demand the productive agents sell their own finished products under competitive market conditions.

If all of these assumptions were realized in practice, the distribution of income in actual economic life would be like that envisaged by the general theory. When the assumed conditions are not present in practice or are only partly realized, the general theory of distribution is subject to a number of qualifications which we have noted in earlier chapters. We have also noted that the various assumptions are seldom, if ever, completely true or completely false as descriptions of actual market conditions in capitalistic economies. The extent to which they are realized in practical situations varies from one time to another and from one market to another in regard to a given agent of production, and varies from one agent of production to another at the same time and in the same market.

While any of the assumptions may be partly invalid at particular times and for particular grades of productive agents, the assumption that the government does not interfere in the distribution of income has been becoming less and less true in erstwhile capitalistic economies in recent years. In the United States, for example, the federal government during World War II was interfering with or determining rents in many areas of the country. It had in effect a law promoting and supporting organizations of workers for purposes of collective bargaining with employers, a law specifying minimum wages and maximum hours for employees of firms operating in interstate commerce, and a law under which it placed "ceilings" on wages and salaries. It was carrying on financial policies which had a considerable effect on interest rates in the economy, and it was profoundly changing the final distribution of money income by means of heavy progressive taxes on incomes and excess profits. In general, considering the validity or invalidity of the assumptions of the theory of distribution as a whole, it is necessary to conclude that the general theory is only a rather rough approximation of the distribution of income which occurs in actual practice in our economic system.

THE PERSONAL OR FAMILY DISTRIBUTION OF INCOME

Whatever we may think about the significance of the general theory of the distribution of income, it is necessary to recognize that it leaves us with incomplete information concerning the status of individual income receivers. The national income is consumed in the last analysis by individuals and families, and the factor which is more important in connection with their scale of living and general economic welfare is the amount of income received per individual or per family *from all sources*, rather than the rates at which income is paid for each unit of land, capital, or labor.

Inequality in Income Distribution. It is quite easy to find the average income per person or per family in a given year by dividing the national income for that year by the number of individuals or families which received the income, but such averages would mean little, for great inequality exists in the distribution of income on a personal or family basis. Many millions of income receivers have low or less than average incomes, while a relatively small number of persons and families have extremely large incomes. The highest incomes received by individuals or families are many thousands of times as large as the incomes received by other individuals or families at the lower end of the income scale, since some individuals and families receive $1,000,000 or more per year while others get less than $500, or even less than $250, per year. This extreme inequality in the distribution of the national income on a personal or family basis has some important economic consequences which will be examined later.

Income Distribution by Families in 1929. Data on the personal or family distribution of income in the United States are not available for all years, but estimates for individual years may be studied. A study of the distribution of income by families in 1929 revealed that more than 21 per cent of all the families in the country had incomes of less than $1,000, more than 42 per cent had incomes of less than $1,500, almost 60 per cent had incomes of less than $2,000, and almost 92 per cent had incomes of less than $5,000. The 8 per cent of all the families, which had incomes of $5,000 or more in 1929, received altogether some 42 per cent of the total income of all families. As shown in Table 17, the highest 10 per cent of the families in 1929

Table 17: A Summary of the Distribution of Income by Families in the United States in 1929

Proportions of Families (*in percentages*)	Proportions of Total Income Received (*in percentages*)
Highest 10	46
Second 10	13
Third 10	9
Fourth 10	7
Fifth 10	6
Sixth 10	6
Seventh 10	5
Eighth 10	4
Ninth 10	3
Lowest 10	1

received 46 per cent of the total income of all families, while the lowest 10 per cent received about 1 per cent of this total income, and the lower 90 per cent received only 54 per cent of the total income. The upper half of

the families in 1929 received 81 per cent of the total income of all families, while the lower half of the families received only 19 per cent of this total income.[1] The income distributed by families in 1929 included some 83 per cent of the total national income, the remainder being received by unattached individuals.

Income Distribution by Families in 1935–1936. An investigation of the distribution of income by families in the United States in the year from July 1, 1935, to June 30, 1936, showed that almost 42 per cent of all the families had incomes of less than $1,000, almost 65 per cent had incomes of less than $1,500, about 79 per cent had incomes of less than $2,000, and more than 97 per cent had incomes of less than $5,000. The families which had incomes of $5,000 or more, or less than 3 per cent of all the families, received about 21 per cent of the total income of all families. These data would suggest, at first glance, that the distribution of income by families was considerably more unequal in 1935–1936 than in 1929. However, it must be remembered that the total national income as measured in money was much lower in 1935–1936 than in 1929, and that any such great decline in the national income or total income distributed by families will bring a larger number and percentage of all families below any fixed amount of income per year, such as $2,000 or $5,000, even if the extent of inequality in the distribution of money income by families remains the same from the one year to the other. The ultimate interest of the individual or family is in real income rather than money income and the decline in real income from 1929 to 1935–1936 was not nearly so great as the decline in money income.

Actually, if we may assume that the results of the 1935–1936 study may be compared directly with those of the 1929 study, the distribution of income by families seems to have been less unequal in the later period than in 1929, as may be seen in Table 18. However, even in 1935–1936, the highest 10 per cent of the families received about 36 per cent of the total income of all families, while the lowest 10 per cent of the families received about 2 per cent of this total income. The highest 20 per cent of the families received 51 per cent of the total income of all families, and the other 49 per cent of the total income was received by the other 80 per cent of the families. Again, the upper half of the families garnered some 78.5 per cent of the total income of all families in 1935–1936, which meant that the lower half of the families could receive only some 21.5 per cent of this total income.[2]

[1] The income data for 1929 were adapted from M. Leven, H. G. Moulton, and C. Warburton, *America's Capacity to Consume*. Washington, D.C.: The Brookings Institution, 1934, Chap. 5.

[2] The income data for 1935–1936 are from the United States National Resources Committee, *Consumer Incomes in the United States, Their Distribution in 1935–36*. Washington, D.C.: 1938.

Table 18: A Summary of the Distribution of Income by Families in the United States in 1935–1936

Proportions of Families (*in percentages*)	Proportions of Total Income Received (*in percentages*)
Highest 10	36.0
Second 10	15.0
Third 10	11.0
Fourth 10	9.5
Fifth 10	7.0
Sixth 10	7.0
Seventh 10	5.0
Eighth 10	4.5
Ninth 10	3.0
Lowest 10	2.0

Personal Income Distribution in 1944. Estimates for the United States in 1944 showed that about 32 per cent of all income recipients had incomes of less than $1,000, 86.5 per cent had incomes of less than $3,000, and 96.5 per cent had incomes of less than $5,000. As shown in Table 19, the highest 10 per cent of the income recipients in 1944 received some 32.5 per cent

Table 19: A Summary of the Personal Distribution of Income in the United States in 1944

Proportions of Income Recipients (*in percentages*)	Proportions of Total Income Received (*in percentages*)
Highest 10	32.5
Second 10	12.5
Third 10	11.0
Fourth 10	11.0
Fifth 10	8.0
Sixth 10	7.5
Seventh 10	6.5
Eighth 10	5.0
Ninth 10	3.5
Lowest 10	2.5

of the total income distributed, while the lowest 10 per cent of the recipients obtained only 2.5 per cent of this total income. The highest 20 per cent of the income recipients had 45 per cent of the total income, and other 55 per cent of the income went to the other 80 per cent of the recipients. Again, some 75 per cent of the total income went to the upper half of the income

recipients, which meant that the lower half of the income recipients could get only 25 per cent of the total income.[3]

Although the figures change to some extent from one year to another, the general picture of income distribution by families or persons, as revealed by these studies and estimates, is quite clear and remains relatively constant from one year to another. That is, in general, the upper 20 per cent of the families or persons will ordinarily receive around half of the total income distributed, leaving the other half of the income for the other four fifths of the families or persons. Moreover, the upper half of the families or persons will usually receive three fourths or more of the total income distributed, leaving one fourth or less of this income for the lower half of the families or persons.

Inequality and the Misguidance of Production. At this point the student may say to himself, "Suppose the national income is divided as unequally as these studies seem to suggest. What of it?" It seems best not to leave such a question unanswered, even though the criticisms of inequality in the distribution of income rest largely on an ethical basis and a discussion of them does not seem to fall appropriately under the principles of economics. The present degree of inequality in income distribution is often condemned, not merely by socialists and communists, from whom such criticisms would certainly be expected, but also by more conservative citizens. Although the criticisms which are advanced are quite numerous, many of them center around the contention that a high degree of inequality in the distribution of income is wasteful.

In the first place, inequality is said to be wasteful in that it results, from the social point of view, in the misguidance or misdirection of production. In strict theory, under capitalism, the production of economic goods is supposed to become adjusted to the needs and desires of consumers through the price mechanism. If consumers had equal incomes, they would probably be willing to pay highest prices for economic goods which they valued most and lowest prices for goods which they valued least. On the other hand, business enterprisers, under the influence of economic motivation, would be attracted to the lines of production whose products commanded high prices in relation to their costs, and would leave or fail to enter lines of production whose products sold for prices which were low in relation to their costs. Moreover, other agents of production would be attracted into the favorable lines of production and would be driven out of the unfavorable lines. The result would be that the goods which were most strongly desired by consumers would be produced in largest quantities, those which were least desired would be produced in smallest

[3] The income data for 1944 are adapted from *Hearings on the Revenue Act of 1943*, United States Senate, Committee on Finance, 78th Congress, 1st Session, p. 849.

amounts, and the production of all goods would be adjusted to the needs and desires of consumers for these various goods.

It has already been suggested in Chapter IV that great inequality in the distribution of income interferes seriously with the working out of this process in practice. Persons with very large incomes may be willing to pay relatively high prices for commodities and services which are unimportant to them and which furnish only a trifling amount of satisfaction, while individuals with very low incomes may be able to pay only very low prices for commodities and services which are vital to their health and welfare. In this situation enterprisers in following the profit motive are likely to produce luxuries and trivialities for the rich while the poor go without adequate quantities of the necessities of life, for enterprisers cannot bother to investigate why prices are high or low, if indeed they are even interested in knowing. For a rich man they will produce an elaborate dinner which costs more than a poor man can pay for food for his family in a whole year. For a rich man's wife they will produce a pair of hose which costs a sum which would furnish a family of average size with a decent standard of living for a year.[4] For a rich man's daughter they will stage a coming-out party which costs more than the aggregate life earnings of some twenty or more poor workers.

Thus we see why rich women may wear coats which cost many thousands of dollars while poor people have no coats at all, and why a rich couple may rattle around in a fifty-room mansion, while eight members of a poor family are stuffed into a couple of squalid rooms. The goods which are produced for the rich do not constitute misdirected production from the point of view of business enterprisers or of the people who consume the goods, but from the point of view of society as a whole the production of these things may be regarded as most inappropriate. In a society in which individual or family incomes were substantially equal, it would be expected that necessities would be produced for all members of society before needless luxuries were produced for any members.

Inequality and the Utility of the National Income. The same general criticism is often expressed in another way by saying that great inequality in the distribution of income prevents us from obtaining the maximum possible aggregate of satisfactions from the consumption of the national income. If we assume rationality of behavior, an individual should spend his first, say, $500 of money income in a given year for those economic goods which he needs or desires most. A second installment of $500 would be spent for other important goods, but these goods would not be as vitally necessary as those secured with the first installment of income, and so on, until the $500 which made a person's annual income $500,000 instead of $499,500

[4] The *Philadelphia Evening Bulletin*, December 5, 1937, noted hose at $2,000 a pair. They were of sheerest chiffon, ornamented with diamonds hung in pendants of platinum.

would probably be spent for something which would give only a very small amount of satisfaction indeed, if in fact a person with such an income could find anything to buy with his last $500 installment of income. Thus even the receipt of money income may be subject to the Law of Diminishing Utility.

It is argued further that, if $500 or some other sum were taken from a person with a very large annual income, it would reduce his total of satisfactions almost infinitesimally, while the same amount added to the income of a person who now receives only $1,000 or $1,500 per year would give him a greatly increased sum total of satisfactions. This does not mean that we should engage in a "share-the-wealth" program, that money income should be taken from the rich and be given directly to the poor, or that the ideal economic system would be one in which all persons or families had precisely equal incomes. Even most socialists do not go nearly that far. The argument suggests only that the consumers of the nation as a whole would derive a greatly increased aggregate of satisfactions from the consumption of the national income if that income were *more equally divided* than at present.

The arguments concerning the effects of inequality on satisfactions in consumption rest on somewhat dubious grounds. That is, they involve interpersonal comparisons of utilities or satisfactions, and there is no way in which such comparisons can be made effectively. Since utility or satisfaction is a subjective affair, one cannot *prove* that rich people enjoy their luxuries less than poor people enjoy their necessities or that one man enjoys his tenth $500 increment of money income less than another enjoys his fifth $500 increment. In fact, it is frequently contended that some people have greater capacity for experiencing satisfactions or gratifications than others. If this is true (and it is not susceptible to proof), a relatively equal division of income would be undesirable from the point of view of maximizing the aggregate satisfactions derived from the national income. However, it seems very doubtful that some persons have capacities for experiencing satisfactions which are thousands of times as great as those of other persons, and differences in income of this magnitude are probably not suited to the purpose of deriving maximum aggregate satisfactions from the national income.

Inequality and the Productivity of Labor. Great inequality in the distribution of income, in addition to any effects which it may have on satisfactions in consumption, is likely to have an important influence on the productivity of the recipients of large and small incomes. At the lower end of the income scale, it leaves great numbers of families with incomes from all sources which are too small to permit the workers to function at full efficiency. As to the amount of money income necessary to maintain the family of average size on a scale of living regarded as a minimum for

"health and decency," estimates vary considerably from one writer to another. They vary because of changes in the price level and purchasing power of money from time to time and from place to place as well as because of differences of opinion between writers as to the items which simply must be included in such a minimum scale of living. However, the estimates often run from $1,500 to $2,000 per year or even more.

If we accepted the low estimate of $1,500 as appropriate for 1935–1936, it would still be necessary to conclude that some 65 per cent, or almost two thirds, of all families in the United States failed to receive a money income in that year which would provide them with such a minimum scale of living for health and decency. If $2,000 per year were taken as the amount of money income necessary for the minimum scale of living, 79 per cent, or almost four fifths, of the families in the United States fell short of the minimum scale in that year. Even in 1944, some 32 per cent of all income recipients had incomes of less than $1,000, and an additional 54.5 per cent with incomes between $1,000 and $3,000 had an average income of only $1,926 after taxes.

The effects of such a situation are easy to describe even though they cannot be measured accurately. When family incomes are below the minimum standard for health and decency, when workers have to get along with inadequate food, clothing, and shelter, when they have to stay on their jobs in spite of illness because they cannot afford to take time off, when they must return to work because of financial considerations even though inadequately recovered from serious illness or accident, when they must get along on a diet of all work and no play or recreation, it seems most unlikely that the efficiency of these workers can be anything like fully maintained. With incomes which are inadequate for present consumption, millions of families must find it impossible to save much for a rainy day, and their efficiency as workers suffers as a result of economic insecurity. They are only too conscious of the devastating effects which they and their families must suffer if they become unemployed because of business conditions, if they are laid up because of illness or injury, or if they lose their jobs because they become too old for industrial or business employment many years before their needs and wants as consumers have ceased. Damocles was greatly disturbed by a single sword hanging over his head. To what extent must workers be affected by this bevy of swords constantly dangling over them! Of course, adequate social insurance of all kinds would go far toward providing for this situation. Although we have made a beginning in some aspects of this wide social problem, few people would argue that our present system of social security is so satisfactory that it fully relieves workers of all their worries connected with unemployment and other matters.

The losses to society which result from the very low incomes of mil-

lions of American families do not end with the lessened efficiency of the workers of the present day. Also to be considered are the effects of low family incomes on the young members of the worker's family. When the children of the poor grow up under unsatisfactory home conditions in an atmosphere of poverty and despair, when they are unable to profit by their educational opportunities because they lack food and clothing, medical attention, or aids to sight and hearing, when they are compelled to go out into the world and take a job at an early age in order to help with the support of their families, society suffers another loss of untold proportions. When the mental and physical development of workers' children is blighted by low family incomes, the potential efficiency of our economic machine also suffers.

The waste which results from inequality is apparently not confined to the lower end of the income scale. Although some persons with large incomes are active and industrious, there are others who really deserve the title of "idle rich" and consume heavily while making no contribution personally to the productive process. They are kind enough to let business enterprisers use their land and capital in return for a generous income, but in many cases this wealth has been received through inheritance and not as a result of their own efforts in the past. It is probable that the amount of productive ability which goes to waste in this way is not extremely large in the aggregate, but most of these persons could probably contribute something to the processes of production if it were impossible for them to lead a life of luxury without working.

Critics of the extremes of wealth and poverty also object to these inequalities because they feel that neither high nor low incomes are entirely deserved by their recipients. No person, whether he receives a high or a low income, is entirely a self-made man, although persons who receive large incomes often flatter themselves by thinking so. Even if he has not profited by inheritance, the man of large income has been helped in reaching his high position by educational opportunities, by free enterprise, by private property, and by other institutions which make it possible for a person of ability to acquire great wealth and income and by many other things which go to make up what is sometimes called "the social heritage." On the other hand, there are comparatively few persons of small income who have not been held back to some extent by environmental conditions which are beyond their control. According to this point of view both large and small incomes are at least in part a social rather than an individual product.

Inequality before the Law. Finally inequality in the distribution of income is criticized because it is said to lead to inequality in other matters. It may be stretching the truth somewhat to suggest that there is one law for the rich and another for the poor, but too often the fate of an individual in either civil or criminal cases is influenced by his wealth and

income status. If a poor man commits a premeditated murder, he is not unlikely to pay the supreme penalty for his offense. If a rich man or his son commits the same crime, he may be able to hire a good lawyer at a high fee who can persuade a jury that the defendant is insane, or was insane at the time of his action, or that he suffered from some psychological disturbance which should relieve him of at least a part of the responsibility for his deed. As a result of such eloquence the murderer may be merely incarcerated temporarily in an asylum, or given life imprisonment or a term of years in the penitentiary. If a poor man steals a few dollars or some groceries because he is out of work and desperate in his inability to provide food for his family, he is likely to be taken out of circulation for several years in order that the majesty of private property may be upheld. On the other hand, a rich man who sells at high prices security issues which later turn out to be worthless or almost worthless, may escape prosecution altogether. Or if he is haled into court, the jury may decide that the poor old fellow did not mean to do any harm or defraud anyone and that he simply made a very natural human error in becoming overoptimistic about the securities which he placed on the market. Eventually this well-meaning person may be retired by his company on a pension.

If a poor man is run over by a careless driver and sustains a broken leg, the motorist's insurance company may magnanimously offer to pay one half of his hospital and doctor bills, and carry it off. If the victim is in comfortable financial circumstances, is able to wait a long time for a settlement if necessary, and is willing and able to fight the matter through the courts while standing all the expenses of legal procedure, he is likely sooner or later to get an adequate settlement in or out of court.

Economic Inequality and Political Inequality. A man of large income is said to have a better chance of getting the government to respond to his wishes than a poor man has of achieving the same result. In theory, the rich man has only one vote, which can be offset by that of any one of his employees, but in practice the rich man may command many votes besides his own. He would not usually be so crude as to try to tell his workers how to vote, but he may suggest, by a slip of paper in the pay envelope or otherwise, that if Mr. Undesired Candidate is elected to be president or governor, business conditions are likely to get worse, the company will probably be forced to retrench, and many workers will have to be laid off. For control purposes such a "word to the wise" may be sufficient.

If a rich man greatly desires the passage or failure of a certain piece of legislation before Congress, he may affect the result by talking like a father to a senator or representative whose nomination or election he has sponsored or influenced. He may send lobbyists to Washington and pay them large sums to work for or against the legislation, or send flocks of

telegrams to senators or representatives asking for the passage or rejection of the bill. By himself, or in cooperation with others of his kind, he may sponsor a press or radio campaign for or against the bill. Finally if, in spite of all his efforts, an undesirable law passes, he can hire able and expensive lawyers to challenge its constitutionality, or to find ways and means by which he can get around it. If a poor man favors or opposes a certain bill which is before Congress, his best course of action would be to go home and pray for its passage or rejection, because there is very little else that he can do. Thus, according to the holders of this point of view, whatever the political theory may be with respect to these matters, inequality in the distribution of income and wealth leads to great inequality in matters of politics and government and helps to prevent the existence of real political democracy.

In Defense of Inequality. In spite of these caustic comments of critics, the present system of great inequality in the distribution of income has its staunch defenders. They say that a considerable degree of inequality is necessary if the economic system is to be efficient and progressive. Efficient production must be indirect, large scale, and specialized, but such production requires large amounts of capital. Capital formation and accumulation depend on saving, and saving in our system can be accomplished only by individuals who have surplus incomes above their current needs for consumption. If our national money income were equally divided among individuals or families, little if any saving might occur. With some persons poor and others rich, we may expect individuals with large incomes to save extensively and devote the saved funds to capital purposes. This line of reasoning also argues that individuals with large incomes perform other valuable services for society, such as devoting large sums of money to the support of art, literature, and music or setting up foundations for medical or educational research. These things too, it is claimed, might be lost to society if the national income were equally or almost equally divided among individuals or families.

These arguments require some examination. In the first place, we must admit that individuals with large incomes in our economic system do perform the functions which have been mentioned and that these things are useful in general. If the national income were so small that, when equally divided, it furnished each individual or family with an amount of money income no greater than was necessary for immediate consumption, it might be that these functions *would not be performed.* However, there are some persons who would argue that these functions, with the possible exception of some saving and capital formation, *should not be carried on* in such a situation. The same small national income, distributed with great inequality, might result in the performance of these functions by the rich at the expense of extreme poverty on the part of millions of

individuals and families at the other end of the income scale. However, if the national income were great enough so that, when rather equally divided, it furnished each individual or family with more than enough money income for immediate consumption needs, it is probable that art, music, and literature would be supported, that desirable research projects would be carried on, and that an adequate amount of saving and capital formation would occur. If private individuals and families did not undertake these projects, the various units of government could sponsor them.

Inequality in income distribution is commonly regarded as necessary in order that adequate incentives may exist for individuals in our economic system. If the national income were divided equally, it is asked, how would we stimulate workers to produce efficiently? How would we get the individual to accept the most difficult and responsible position in economic life for which he can qualify? How would we get people to be business enterprisers and managers if they received about the same income as ordinary workers? Surely, if our economic system is to be efficient and progressive, it is necessary to hold out to the individual the chance to better his economic position and we must provide unusual economic rewards for unusual accomplishments.

It must be admitted that there is no final answer to these questions. They assume, of course, that the individual is economically motivated, but the same assumption is made by economists on the ground that it is more nearly true of people in general than any other broad statement which could be made concerning motivation. However, the economist does not deny that there are many other motives which are of some significance in economic activity, and no one knows just to what extent these other noneconomic incentives could be developed and made effective if the national income were equally or almost equally divided. Is the individual acquisitive because he lives in an acquisitive society called capitalism, or do we operate an acquisitive society because the individual is fundamentally and naturally acquisitive? Whole volumes would not answer this question to everyone's satisfaction. However, even if economic motivation were a part of human nature and some differences in income were necessary as incentives to efficiency and progress, it might still be suggested that adequate incentives could be furnished to the individual without having the national income divided nearly as unequally as it is at present.

Why Does Inequality Exist? The original basis for unequal incomes among individuals and families is found in income derived from the sale of labor services. The labor supply is divided into noncompeting groups in such a way that many persons can perform the functions required by occupations in the lower groups, while comparatively few persons can fill the requirements of the occupations in the higher groups. Significant differences in the marginal productivity of units of labor service therefore

exist between the labor groups, and, although wages do not actually adjust themselves accurately to these intergroup differences in marginal productivity, very large differences in wages or payments for labor service are found between workers in different groups. Even if there were no other sources of income for various individuals, a considerable degree of inequality in income distribution would result from this factor alone, and thus inequality would be self-perpetuating to a great degree in view of the effects of low incomes on the efficiency and productivity of many workers and their children.

In practice, since our economic system permits private individuals to own land and capital and to entrust these objects of wealth to various enterprisers or lines of production on the basis of the income to be derived from their use in production, individuals receive income for the use of their land and capital as well as for the use of their labor. This fact might make for equality rather than inequality in the distribution of income if those individuals who had small labor incomes derived large incomes from owning land and capital while the individuals who had large incomes from the sale of labor services received little if any income from the ownership of land and capital. However, the actual situation is just the other way around. Reliable estimates of the distribution of wealth in the United States are even more difficult to obtain than those for the distribution of income, but the studies which have been made indicate that ownership of wealth is much more concentrated than that of income. Individuals and families must have some income in order to live, but they may receive fairly large annual incomes without ever coming to own any appreciable quantity of wealth.

It is to be expected, therefore, that individuals who receive small labor incomes will not own any significant amount of land and capital or receive large amounts of rent and interest to bolster up their labor incomes, while individuals who receive large incomes from the sale of labor services will be likely to own large amounts of land and capital and receive large additional incomes in the form of rent, interest, and sometimes profits. Inequality in the distribution of income would be great enough if individuals were limited to receiving income from the sale of their labor services and from the wealth which they could accumulate in their own lifetimes, but it becomes still greater when the institution of inheritance is allowed to operate. Individuals who amass great fortunes are permitted to pass them on to their heirs, who may continue to receive income from the inherited wealth in addition to that which they derive from the sale of labor services and from wealth which they can themselves accumulate.

Fundamentally, then, inequality in income distribution is derived from the institutions of capitalism, such as private property (including inheritance), free enterprise, competition, and the pricing process. Under these

institutions, the relatively scarcer and more productive agents of production come to command high rates of remuneration while the less productive and relatively more plentiful agents of production command low rates of remuneration. This result may be considered desirable in many respects. High rates of remuneration for the scarcer agents tend to reserve these agents for the uses which are deemed most important on a price basis and keep them from being wasted in uses in which they would have only a low value-product. Low rates of remuneration for the relatively more plentiful agents of production tend to lead to their use in large quantities and to prevent their being wasted in unemployment as they would be if high rates of remuneration were set upon them. High rates of remuneration for the scarcer agents tend to bring about an increase in their quantity, where such a result is possible, while low rates of remuneration for the relatively more plentiful agents tend to produce a restriction of their quantity. However, under the institutions of our economic system, high rates of remuneration for the scarcer agents of production mean large incomes for the owners of these agents, while low rates of remuneration for the relatively more plentiful agents give small incomes to the owners of these agents. The result is inequality in the distribution of income among persons and families.

Remedies for Inequality. It should be noted that the actual extent of inequality in the distribution of income in the United States is not so great as the data presented earlier in the chapter seem to suggest. The studies of income distribution by individuals or families show how income was distributed before taxes were paid. Among other taxes we have a federal income tax which provides high rates of taxation in the upper income brackets and which serves, in combination with other taxes, to reduce the very large incomes to a fraction of their original size. However, inequality in the distribution of income remains very great even after this factor has been taken into consideration. It is also true that the differences in money incomes received by individuals and families do not accurately represent differences in consumption and scales of living, since individuals who receive large incomes often save a considerable part of these incomes.

Inequality in income distribution in the United States is great enough so that many people think that it should be reduced considerably. The proposals for this purpose are very numerous. Inheritance taxation could be extended, land rents might be taken away by means of a single tax or other device, income and excess profits taxes could be maintained at high levels which have heretofore been known only in wartime, governmental attempts might be made to break down the environmental barriers to movement between labor groups, a great variety of free services or even commodities might be furnished to the poor, and so on. However, in discussing the pros and cons of inequality we have already wandered away from the

area previously marked out as appropriate for discussion in economic theory and we cannot go further afield to deal with the merits and demerits of various schemes for reducing inequality in income distribution. Students who would like to pursue this subject further are referred to any standard work in the field of economic problems.

NONCAPITALISTIC DISTRIBUTION

Distribution under Socialism. Although there is much dispute as to what could be or should be done about inequality in income distribution under capitalism, there is little doubt as to what the socialists would do about it. Under the ideal system visualized by socialists, land and capital would be owned by society as a whole and not by private persons, while the various lines of production, or almost all of them, would be operated by society as a whole through various governmental units. In this way, rent, interest, and profits would be eliminated as sources of private income, and the national income, or such part of it as was made available for consumption, would be distributed to individuals in the form of wages. There are some socialists, of whom George Bernard Shaw has been a leading example, who believe that the national income should be divided among persons with absolute equality, but most socialists would permit moderate differentials in wages to exist under socialism. Land and capital might be arbitrarily assigned to different lines of production by the economic planners, but it would be neither efficient nor humane to accord workers the same treatment. For this reason wage differentials between industries and occupations would be set up for the purpose of inducing workers to distribute themselves among occupations and industries in a manner deemed appropriate by the economic planners.

How large the wage differentials would be under socialism cannot be definitely predicted, but it is certain that they would be much smaller than those which prevail under capitalism. Moreover, the wage differentials would be used not only to reward the workers for their varying accomplishments in production but also as a tool to induce the workers to do the things that needed to be done. It should be obvious that our theories of distribution would be worthless in connection with a socialistic economic system. Descriptions of the determination of rent and interest in terms of marginal productivity or other factors would not be helpful in studying an economic system in which individuals are not allowed to receive these returns as private income or own the agents which produce the returns. Even the theory of the determination of wages in terms of marginal productivity would seem to have relatively little application to the determination of wages under socialism.

Our discussion does not mean that the factors which give rise to rent, interest, and profits under capitalism would all be eliminated under

socialism. Land would still be of various grades and qualities under socialism, and a given amount of capital and labor applied to high-grade land would still be more productive than if it were used on low-grade land. However, the differential surplus, instead of resulting in rent for individual owners of high-grade land, would merely belong to society as a whole, and the workers of the system in general would receive the surplus as part of their real wages. In similar fashion, under socialism as under capitalism, production via the indirect process using large amounts of capital would be more efficient than direct production, and saving would be necessary in order that capital formation might take place. However, the saving would be done by all members of society under socialism and would be accomplished by their going without consumers' goods in the present; the increased quantities of economic goods which roundabout production would make available would then be shared by all members of society in the form of increased real wages.

Finally, the conditions which give rise to profits could hardly be completely eliminated under socialism. While some risks might be eliminated or minimized under socialism, weather and other natural conditions would still remain uncontrollable, and the actual efficiency of workers might be much greater or less than had been planned. However, any failures of production to come up to expectations, or any successes in getting production to exceed planned estimates, would be of concern to the members of society as a whole and not to a particular class of persons called business enterprisers. Thus the production principles which are involved in the theory of distribution would also be valid under socialism, but our description of functional and personal distribution of income would fall by the wayside. The Soviet Russian economic system is not one of pure socialism, but the characteristics of the process of income distribution in Soviet Russia are roughly those which have been described for theoretical socialistic systems.

Distribution under Communism. Under communism, consumers' goods as well as land and capital would probably be owned by society as a whole. Since the use of money would apparently be abandoned, commodities and services would have no money prices and individual workers and consumers would have no money incomes. All individuals who were able would work for a living, but the income received for productive services would immediately take the form of real income. It is the communist notion that individuals should contribute to production according to their abilities and receive real income according to their needs. This is sometimes taken to mean perfect equality in income distribution, but this conclusion is not justified unless one is willing to assume that all individuals or families have precisely the same needs in consumption. As under socialism, those who produced the national income would share in the products turned out by all the agents of production working together, but it would be point-

less to try to trace these products back and allocate them to the functioning of individual agents of production. It is hardly necessary to say that our analysis of the determination of rent, interest, wages, and profits, and our description of the personal distribution of income, would be meaningless in connection with a communistic economic system.

Distribution under Fascism. Under fascism the institutions of capitalism prevailed at least in name. Land and capital were owned by private individuals for the most part, many industries were privately operated, economic activity was at least partly competitive, and all the familiar shares in distribution which we know under capitalism were present. On this basis, it might have been thought that our theories of distribution would be as valid for fascism as for capitalism, but this conclusion was not really warranted. In the fascist countries, such as Italy and Germany, the leaders of the government were free to interfere with and control economic activity to any desired extent and in practice did interfere to an overwhelming extent, especially in wartime. They were quite capable of ordering business enterprises to invest all earnings, above a certain rather meager rate, in government bonds, of putting prohibitive taxes on dividend payments so as to force businesses to reinvest their earnings, of ordering general reductions in wages and prices throughout the economic system, of limiting the rents which could be charged, or of placing burdensome capital levies on the value of land or the capital and surplus of corporations.

Thus while rent, interest, wages, and profits existed as shares of income for private individuals under fascism, the only principle that could be advanced in connection with these distributive shares was that they would be whatever the government and party leaders wanted them to be, which is the same as having no principle at all. Even if the leaders had allowed a large part (say half or more) of the fields of economic activity to operate without direct interference, our theories of distribution would not have been valid even for this unrestricted section of the economy. That is, it would not have been possible to make the distributive shares in this free section of the economy entirely independent of the conditions which prevailed in the controlled section of the economy. The conclusion is, therefore, that the extent of governmental interference with economic activity under fascism was always too great to permit our theory of distribution—which depends, among other things, on the markets for the productive agents being free of governmental interference and control—to be even roughly applicable in these economic systems. Great inequality in the distribution of income by individuals and families existed under fascism, but there as elsewhere its extent was modified by heavy taxation.

QUESTIONS AND PROBLEMS

1. "The discussion of the distribution of income in terms of the rates at which the owners of the productive agents are paid overlooks one of the most important questions in connection with the distribution of the national income." Explain.

2. "The most striking feature of the distribution of income by families or persons in the United States is found in extreme inequality." Discuss.

3. "Data on the personal or family distribution of income in the United States for 1929, 1935–1936, and 1944 suggest that inequality is gradually decreasing." Show whether you agree.

4. Why is it argued that great inequality in the distribution of income results in the misguidance of production? Explain.

5. How does extreme inequality in the distribution of income prevent the realization of the greatest possible aggregate of satisfactions from the consumption of the national income? Explain.

6. "Great inequality in the distribution of income decreases the efficiency of production." Do you agree? Explain.

7. "Inequality in the distribution of income leads to inequality between persons with respect to other matters." Show whether you agree.

8. Could a capitalistic economic system be operated without a considerable degree of inequality in the distribution of income? Explain.

9. "Considerable inequality in the distribution of income is necessary in order that there may be adequate saving and capital formation under capitalism." Show whether you agree.

10. "If our economic system is to be efficient and progressive, it is necessary to provide unusual economic rewards for the unusual accomplishments of individuals. Hence, great inequality in the distribution of income is necessary and desirable." Do you agree? Explain.

11. Why does great inequality in the distribution of income exist in our capitalistic system? Explain.

12. "Inequality in the distribution of income results in part, but only in part, from the sale of labor services of unequal market value." Explain.

13. "Though the result is inequality in the distribution of income, it is economically desirable that high prices be set on the scarcest agents of production and low prices on the more plentiful agents." Explain.

14. "Inequality in the distribution of income would be much less under socialism than under capitalism." Why?

15. What do the socialists mean when they say that rent, interest, and profits would be eliminated under socialism? Explain.

16. Compare socialism with communism with respect to the distribution of income among persons.

17. "There were really no economic principles which controlled the distribution of income among persons under fascism." Show whether you agree.

REFERENCES FOR FURTHER READING

Boulding, K. G., *Economic Analysis*. New York: Harper & Brothers, 1941, Chaps. 11, 12, 14.

Brown, H. G., *Basic Principles of Economics*. Columbia, Mo.: Lucas Brothers, 1942, Chaps. 12–16.

Bye, G. R., *Developments and Issues in the Theory of Rent*. New York: Columbia University Press, 1940, Chaps. 1–5.

Bye, R. T., and Blodgett, R. H., *Getting and Earning*. New York: F. S. Crofts & Co., 1937, Chaps. 1–8.

Carver, T. N., *The Distribution of Wealth*. New York: The Macmillan Company, 1913, Chaps. 4–7.

Clark, J. B., *The Distribution of Wealth*. New York: The Macmillan Company, 1899, Chaps. 9, 12–13.

Henderson, H. D., *Supply and Demand*. New York: Harcourt, Brace and Company, 1922, Chaps. 5–10.

Hicks, J. R., *Theory of Wages*. London: Macmillan & Company, Ltd., 1935, Chaps. 1–6.

Knight, F. H., *Risk, Uncertainty, and Profit*. Boston: Houghton Mifflin Company, 1921, Chaps. 2, 9–10.

Marshall, Alfred, *Principles of Economics*. London: Macmillan & Company, Ltd., 8th ed., 1920, Book V, Chaps. 8–10; Book VI, Chaps. 1–11.

Stigler, G. J., *The Theory of Competitive Price*. New York: The Macmillan Company, 1942, Chap. 10.

XXIII

Money

Our discussions of production, price determination, and the distribution of income have suggested that a tremendous volume of exchanges must take place in an economic system in which production is indirect, large scale, and specialized. Because the possibilities of direct exchange or barter are limited and the difficulties involved in it are considerable, the modern economic system depends very largely on the process of indirect exchange, which requires the use of money or credit or both as a medium of exchange. Therefore the next step in the study of our economic system is an analysis of money, which will be followed in turn by an investigation of credit and a discussion of banking institutions.

THE NATURE AND FUNCTIONS OF MONEY

The Definition of Money. For many years money was commonly defined as any article which serves as a standard of value and which because of general acceptability is used primarily as a medium of exchange. However, a definition which requires that an article be both a medium of exchange and a standard of value in order to be money is not very satisfactory in view of the present monetary situation in the United States. At present our standard monetary unit, the dollar, is defined by law as $15\ 5/21$ grains of gold nine tenths fine or pure. The dollar as so defined is our standard of value for things in general. In spite of these facts, gold and gold coins do not circulate in exchange, the gold stock of the country is all held in the form of bullion by agencies of the federal government, and the possession of gold by private individuals or firms, except for purposes of manufacture, export, or personal wear, is against the law of the land.

Since no actual gold dollars are coined or circulated, all the kinds of money which pass as dollars (or multiples and fractions of dollars) and which serve as a medium of exchange are really only substitutes for actual dollars. In other words, the things which are used as a medium of exchange are not the same as that which forms the basis of the monetary system and serves as the standard of value. In view of this situation it seems desirable

to define money as *any article which serves as a standard of value, or is generally acceptable and used primarily as a medium of exchange, or both.* However, if the articles used as a medium of exchange are not the same as that which functions as the standard of value, ordinarily they will not serve very satisfactorily for long as a medium of exchange unless some fairly close relationship exists between them and the article which constitutes the standard of value of things in general. In this country the things which serve as the medium of exchange are related directly or indirectly to reserves of gold, which is the article that serves as the standard of value, even though the gold may not be used as the medium of exchange.

The definition of money is stated in general terms; before we go further it is necessary to give more specific content to the definition. Some writers interpret the definition to include both common money (such as bullion, coins, government paper money, and bank notes guaranteed by the government) and bank money, or checks drawn by individuals and firms against demand deposits in banks. However, we shall use the term money to mean only bullion, coins, government paper money, and bank notes guaranteed by the government. We exclude checks, drafts, promissory notes, and securities from the category of money on the ground that, although they may function as a medium of exchange in individual transactions, they lack something of the quality of general acceptability which is required of money.

If an individual, in the course of his travels, makes a purchase in a store in a city far from his home, he will have little difficulty in making his payment in any form of money in circulation, but he may run into a good deal of trouble in attempting to use his personal check for the same purpose. The storekeeper may well wonder whether his customer is really the person whose name is signed to the check, whether he has an account in the bank on which the check is drawn, whether there is a large enough balance in the account to cover the check, and so on. It would be even more difficult to induce the storekeeper to accept a promissory note, by which the customer promised to pay him the value of the goods after sixty days or some other period of time, or to draw a draft, or order to pay, on his customer. Checks, notes, drafts, acceptances, and other credit instruments are used in financing a tremendous number of exchange transactions. In fact, the transactions in which actual money functions as a medium of exchange are generally thought to be a small fraction (by value) of the entire volume of exchange transactions in our economy. However, the use of checks, notes, drafts, and acceptances requires special knowledge and conditions which are not essential to the use of ordinary money; hence it seems desirable that we distinguish between money and credit instruments on the basis of general acceptability, while using the term "money and

credit" to refer to all the things which may be employed as the medium of exchange.

The Functions of Money. The discussion of the meaning of money has suggested some of the functions of money but we shall now state these functions more specifically. In the first place, money, or part of it at any rate, serves as a medium of exchange. Instead of exchanging commodities and services directly for each other, we sell commodities and services for money and use the purchasing power thus acquired to buy other commodities and services which we desire. Money is readily divisible for making small purchases and, because of its general acceptability, it is usually possible to find a seller who will accept money in exchange for any commodity or service which we desire, or to find someone who will offer money for the commodity or service which we wish to sell. Thus the use of money enables us to get around one of the chief difficulties of barter: the difficulty experienced in finding some person who not only wishes to acquire what we have to sell but desires to dispose of the thing which we wish to buy. Although credit instruments as a whole are used as a medium of exchange to a greater extent than money itself, these money substitutes, as we shall see, are themselves dependent upon the money of the country.

Second, money serves as a standard of value, or means of expressing and comparing the values of things in general. Quite apart from any desire to exchange commodities and services, we may say that a certain watch is worth $35, a pair of shoes $10, or a share of stock $125. Money does not give us a completely reliable and unchanging measure of the value of other things, because the value of the unit of measurement itself changes. If a child's height measures thirty inches at one time and forty-five inches at another, we know that the change is due to the growth of the child and not to a change in the size of the inch. But a change in the price of a suit of clothes from thirty to forty-five dollars might be due to an increase in the value of the suit or to a decrease in the value of the unit of measurement, the dollar. It is really more accurate, then, to say that money is a common denominator of value than to say that it is a standard of value. Nevertheless, it is useful and convenient to express the values of all sorts of things in terms of money, and to use money as the unit of account. Business accounts and records are kept in terms of money, and such complex things as national wealth, income, production, and foreign trade are usually measured in the same way.

Again, money functions as a store of value. Money may be of great assistance when individuals must sell perishable commodities or services at present, or even when they wish to sell nonperishable commodities now in order to avoid the fuss and bother of storing them, and in neither case wish to obtain other commodities and services in exchange at the time. By making the present sale for money, and keeping the money until commodi-

ties and services are desired in return, the individuals are able to store values to an extent which would be quite impossible under barter. Money does not serve perfectly as a store of value because its own value is likely to change through time, but it serves the purpose better than most things, since whatever is used as the basis of the monetary system is usually chosen in part because of its relative stability of value.

Finally, money is the basis for credit and credit instruments. When the payment for goods purchased is deferred until a later time, promissory notes or drafts are drawn in terms of money and call for payments of money. Borrowing and lending transactions are carried on in terms of money, statements of debt or ownership in connection with bonds or stocks are expressed in similar fashion, and checks drawn on demand deposits in the banks are orders for the payment of money. The whole system of credit and credit instruments rests upon a monetary base, while money also functions as reserves behind certain forms of credit such as bank deposits.

The Characteristics of Money. Many things have been used as money at one time or another in history. Besides the precious metals, gold and silver, such things as iron, salt, tobacco, ivory, cattle, wampum, furs, and arrowheads have been used in particular regions and times. Most of these things lack one or more of the qualities which enable an article to serve effectively as money. In the first place, an article which is to be used as a medium of exchange must be generally acceptable, but such general acceptability depends to a great extent upon the possession of other characteristics. It should combine large value with small bulk so that it is portable, it should be durable enough to resist wear and intentional defacement or clipping, it should be uniform so that each piece is like every other piece which has the same value, it should be divisible without loss of value, and it should be cognizable so that anyone can easily decide the value of a particular specimen. The article which serves as the standard of value should ordinarily be valuable in other than monetary uses, and, if it is to function successfully as a store of value and basis of deferred payments or credit, it should have relative stability of value in terms of other commodities and services.

The articles which most nearly possess the characteristics of a good money commodity are gold and silver, and especially gold, but these articles are not perfect by any means. Gold has large value in small bulk and is readily portable, but silver does not qualify as well in this respect, since, for example, a thousand silver dollars would make a sizable burden. Gold and silver are divisible without loss of value, but gold coins of small denominations, such as dollars, half dollars, or quarter dollars, would be too small for practical use. Gold and silver, when coined, do rather well with respect to durability, uniformity, and cognizability, but their stability of value is not entirely satisfactory.

Paper money is commonly used as a substitute for metallic money and

does not need to have all the characteristics of a good money commodity. However, it is portable, uniform, and cognizable, and it can be made in any desired denomination. On the other hand, paper money wears out rather easily, and its stability of value is no greater or less than that of the metallic money on which it ordinarily depends. When paper money is issued for its own sake and without reserves of metallic money behind it, its instability of value is notorious because it is usually overissued.

COINAGE AND CURRENCY

Full-bodied and Fiduciary Money. We have seen that the money of a country is ordinarily of two kinds—metallic and paper money—and we now turn to questions of the coinage and use of metallic money or coins. Metallic money may be either full-bodied (standard) or fiduciary. As we shall see later, paper money is always fiduciary in character. By "full-bodied money" we mean money which has an intrinsic value equal to its face value or, in other words, which has a value as a commodity equal to its value as money. For example, the gold coins used in the United States before 1933 were full-bodied money. If a ten-dollar gold piece was melted down, the gold bullion derived from it was worth precisely ten dollars as metal. However, this was not true of any coins save gold coins. The silver contained in the silver dollar, half dollar, quarter dollar, or dime was worth much less than these respective amounts as metal. There was not five cents' worth of nickel and copper in the "nickel" or a penny's worth of copper, tin, and zinc in the "cent."

These silver dollars and subsidiary coins were fiduciary money. Though they were worth less as commodities than as money, they served effectively as money. Fiduciary money is so called because its use rests upon the faith or belief of the users that this money can be exchanged for full-bodied money by applying to the government or central bank, or that the money can be used to secure commodities and services or pay debts even if it cannot be exchanged for full-bodied money. The faith in the money may rest upon laws which declare the money to be acceptable in payment of all kinds of debts or which require that all forms of money be kept at a parity of value by making them convertible into each other, or the government may promise specifically to redeem the fiduciary money with full-bodied money upon request.

Coinage. Although it might be possible for metallic money to circulate by weight it is ordinarily used by count or piece. This practice requires that the metal be coined into pieces which are uniform as to size, content, and pattern, and that the coins be manufactured in such a way as to protect the metal from wear and alteration in so far as this is possible. That is, they are stamped with designs on both sides, and the edges of all coins (save the nickel and penny in this country) are raised and "milled" to prevent clip-

ping or shaving. Coinage, then, refers to the making of uniform coins or pieces of the monetary metals, and stamping them in such a way as to guarantee their weight and fineness. The business of coinage might be entrusted to private individuals or companies, but in practice it is universally a governmental monopoly. Having gone to considerable expense to provide a country with adequate coinage facilities and to considerable difficulty to prevent alteration and counterfeiting, the government does not regard competition in coinage as the life of the trade. In fact, the government is likely to resent in the extreme any attempt by private persons to offer competition in this field, and it usually removes convicted counterfeiters from circulation for extended periods of time. Since the engraving and printing of paper money is also a monopoly of the government, it takes precautions in the form of special paper and intricate designs to prevent counterfeiting.

Free and Limited Coinage. The coinage system of a country will be either free or limited. Free coinage needs to be distinguished from gratuitous coinage, even though free and gratuitous usually mean about the same thing. Coinage is gratuitous if no charge is made to the owner of bullion for the service, while it is free if the government stands ready to accept for coinage and make into money any amount of a specified metal which bullion owners may bring in for the purpose. The coinage is limited if the government will take metal for coinage only as needed or only up to certain specified amounts. Coinage may be free even though a charge is made for the service, and it may be gratuitous though strictly limited in amount.

Legal Tender. Another function performed by governments in connection with money is to decide and provide by law which coins and types of paper money are legal tender for private debts and in the relations between governments and citizens. Legal tender is money which must be accepted by a creditor in payment of a debt, unless a specific agreement to the contrary has been made, and which the government agrees to accept in payment of taxes and fines. If one man borrows a hundred dollars from another and they agree that the debt is to be paid with a specific kind of money (say silver dollars), such an agreement is enforcible whether or not silver dollars are legal tender, and the offer of some other kind of money does not suffice even though it is legal tender. If the debt agreement does not specify any particular kind of money, the debtor may offer any legal tender and the creditor must accept this money sooner or later or receive nothing. Refusal to accept legal tender does not cancel the debt, but under the statute of limitations the debt will expire sooner or later unless the creditor decides to accept legal tender. Moreover, interest on the debt stops as soon as the debtor offers legal tender to the creditor in settlement of it. In the United States the situation with regard to legal tender has been greatly simplified since 1933, when all types of metallic and paper money

except gold and gold certificates were made full legal tender for all debts, public and private. Formerly the question of legal tender had been a very complicated one in our monetary system.

PAPER MONEY

Most monetary systems include considerable quantities of paper money, in addition to metallic money. The paper money includes direct obligations of the government and those of central banks or other banks as guaranteed by the government. As fiduciary money, it circulates because of the faith of the users of the money that it can be exchanged for full-bodied money either at once or at some later time, or that, regardless of convertibility, it can be used satisfactorily for all money purposes. Paper money is more convenient than metallic money, especially if individuals need to carry considerable sums around with them. Its use saves the loss which would result from wear in the repeated handling of metallic money. Its quantity can be sharply and quickly increased or decreased, if reserve requirements are flexible, while the total quantities of gold and silver available for use can change only very slowly. Finally, the use of paper money may save expense for the government and taxpayers. If a reserve of 25 per cent in gold is provided behind paper money, a billion dollars' worth of gold will support the issue of four billion dollars in paper currency, whereas four billion dollars in gold would be required if all the money were to be full-bodied.

Representative Paper Money. There are three major types of paper money. Representative paper money is that which has a 100 per cent metallic reserve behind it, held by the government and payable to the holder of the paper money on demand. The metallic reserve may consist of full-bodied money or money which is itself fiduciary. In the case of the gold certificates which circulated in the United States until 1933, the 100 per cent metallic reserve consisted of full-bodied money (gold), but silver certificates are backed by silver dollars or bullion. The reserve behind a one-dollar silver certificate, in other words, is not a dollar's worth of silver but one silver dollar, or enough silver to make a silver dollar, and, of course, there is not nearly a dollar's worth of silver in a silver dollar. There is no question of the government's ability to redeem representative paper money with metallic money, but the government may not be willing to redeem it, as many holders of gold certificates in the United States found out to their sorrow in 1933. Hence representative paper money is fiduciary money and depends upon faith in the government's promise to convert it into metallic money.

Convertible Paper Money. Convertible paper money is that which can be exchanged for full-bodied money upon demand. It is ordinarily backed up by a fractional reserve of full-bodied money as a guarantee of redemp-

tion. The reserve must be less than 100 per cent, or the paper money would be representative, and a specific type of convertible money may have no reserves of full-bodied money at all behind it. Thus silver certificates in the United States have a reserve of silver dollars or bullion behind them, but until 1933 they could be converted into gold (full-bodied money) just as readily as any other kind of money. Under ordinary conditions a fairly small reserve of full-bodied money will suffice as a guarantee of the redemption of convertible money because it is unlikely that any considerable part of this money outstanding will be presented for redemption at any particular time. In fact, people very seldom attempt to convert paper money into full-bodied money unless they suspect or fear, for some reason or other, that such conversion cannot be accomplished. Obviously, on the basis of a given amount of full-bodied money, a much greater quantity of convertible than of representative paper money can be issued. None of the money in actual use in the United States at the present time is convertible.

Inconvertible Paper Money. The third variety of paper money is usually presented as a horrible example of monetary mismanagement. Inconvertible paper money is simply money which cannot be converted into full-bodied money. It is not necessarily paper money which has no reserve of full-bodied money behind it. In the past it has been too often true that governments have issued large amounts of inconvertible paper money because they were in such desperate straits that it was impossible for them to maintain any more desirable type of money in circulation. In such cases the inconvertible paper money has had little, if any, metallic reserve behind it. However, in the United States at the present time the government is required by law to keep large gold reserves behind various kinds of paper money and does in fact hold such reserves, but all kinds of paper money are nevertheless inconvertible. It is quite possible that paper money may have a 100 per cent reserve of full-bodied money behind it and still be inconvertible.

When the government is required to maintain reserves of full-bodied money behind its paper money, and the quantity of paper money which can be issued is a definite function of the amount of full-bodied money which is held in the treasury, the paper money of the country may keep its value very well in spite of being completely inconvertible, and the dangers often encountered in the use of inconvertible paper money are not likely to be important. When reserves of full-bodied money are not held behind the inconvertible paper money or when the quantity of this paper money which can be issued bears no definite relationship to the reserves of full-bodied money, if any, in the treasury, there is grave danger that the inconvertible paper money will be overissued. This is most likely to occur in time of war (or other national emergency), when sufficient funds cannot be obtained through taxation or borrowing to cover necessary governmental

expenditures, and when reserves of full-bodied money have been sharply reduced because of the excess of imports over exports which is likely to occur in wartime.

In such cases it may be impossible to curtail expenditures, and the only alternative seems to be merely to print the paper money with which these expenditures may be carried on. The paper money may carry the promise of the government to redeem it, although the government clearly lacks the power to do so at the time, and it is likely to be made full legal tender for the payment of debts. It may not be difficult to make people use the inconvertible paper money as a medium of exchange, but its value or purchasing power in terms of commodities and services will fall as more and more of the money is pumped into circulation by the government in making its expenditures. Moreover, the process is a cumulative one: the prices which the government must pay for commodities and services rise along with other prices, the necessary expenditures of the government increase, it must print more and more paper money in order to make these expenditures, the prices of commodities and services rise still further, the necessary expenditures of the government increase once more, and so on.

Once the process of inflating the currency is begun, a country is fortunate if the inflation is not carried to disastrous lengths. As prices rise rapidly and considerably and the purchasing power of money falls, creditors are placed in a most difficult position, for the payments of interest and principal which they receive will buy very little. Values of debts and contracts are scaled down or virtually wiped out, persons living on fixed money incomes suffer as the purchasing power of their money incomes declines, and wage earners are affected adversely for, although wages will rise, in a period of inflation it is usually impossible to keep them abreast of other prices. To be sure, production may be stimulated by rising prices for a time, but, if prices go on rising rapidly, business enterprises eventually hesitate to make any productive commitments whatever, and economic stagnation and chaos result.

The Value of Money. It is clear that paper money is more likely to keep a fairly stable purchasing power if reserves of full-bodied money are held behind it than if they are not. However, it cannot be said that such reserves give value to paper money in any direct and automatic fashion. Some people think that, if all money were gold or were backed by a 100 per cent gold reserve, its value or purchasing power would be perfectly stable. This is simply not true. A great increase in the quantity of paper money in circulation, even if backed by a 100 per cent gold reserve, would bring about a considerable rise in prices and decline in the value of the monetary unit, other things being equal. Similarly, if gold money were in actual circulation, a large increase in the quantity of this money would produce a decline in its value and a rise in general prices, if other things remained un-

changed. On the other hand, if it were not overissued, inconvertible paper money backed by absolutely no reserves of full-bodied money might be just as useful and valuable as a medium of exchange as any other kind of money.

The conclusion to which all this leads is often known as the Quantity Theory of Money. It states that, if other things remain unchanged, the value of money varies inversely with the quantity of money in circulation, that is, in actual use in exchange. This means, of course, that the value or purchasing power of money will fall as the quantity of money in circulation increases, and will rise as the quantity of money in circulation decreases, if other things remain unchanged. The factors which are supposed to remain unchanged are the velocity of circulation, or rate of turnover of money, and the volume of trade, or the total quantity of commodities and services which must be exchanged through the use of money.

The assumption that these factors remain unchanged implies that they may change in actual practice and that, when they change, they too may affect the value or purchasing power of money. Thus it may be said that the value of money tends to vary inversely with the velocity of circulation of money, if the quantity of money in circulation and the volume of trade remain unchanged, and that the value of money tends to vary directly with the volume of trade, if the quantity of money in circulation and its velocity of circulation remain unchanged. Clearly a given quantity of money circulating twice as fast as formerly will have the same effect on the value of money as a doubled quantity of money circulating at the former rate, if the volume of trade remains unchanged. And a doubled volume of trade, being exchanged by the same quantity of money circulating at a given rate, tends to affect the value of money in the same way as an unchanged volume of trade which is being exchanged by means of half as much money as formerly, circulating at the given rate. However, under any given velocity of circulation of money and volume of trade, the value or purchasing power of money varies inversely with its quantity.

THE GOLD STANDARD

The Nature of the Gold Standard. The discussion of money up to this point has dealt largely with the nature of money in general and of particular varieties of money. We now turn to an analysis of general monetary systems or standards and consider first the gold standard, which prevailed in the United States for many years. The gold standard means different things to different economists but for our purposes it may be said to include the following essential conditions:

1) The standard monetary unit of the country (such as the dollar in the United States), its weight, fineness, and denominations in which it is issued, are defined in terms of gold.

2) The gold monetary unit is full legal tender in payment of all obligations, public and private.

3) The government maintains a free market for gold; that is, it stands ready to buy or sell gold at a fixed price per ounce or other unit.

4) The government operates a system of free coinage for gold and stands ready to convert into gold coins any quantity of gold bullion which is offered to it.

5) All obligations, public or private, domestic or international, are payable in or convertible into gold. All forms of money can be redeemed in gold, debts and other contracts can call for payments in gold, gold coins may be used as freely as the people desire, and gold can be exported or imported freely in settlement of international obligations.

Advantages of the Gold Standard. The widespread adoption and use of the gold standard in the past rested upon several advantages. In the first place, it was thought that the value of a country's monetary unit could be kept relatively stable by basing it on gold. Gold was demanded not only for monetary purposes but in a number of other uses, so that the total demand was supposed to be highly elastic and capable of absorbing increases in the production of gold without great changes occurring in its value. On the supply side, the annual production of gold was small relative to the total stock already in existence, and production was supposed to adjust itself rather automatically to general business conditions. With the price of gold fixed by the government, changes in production, in the absence of technological changes, depended upon the ups and downs in the cost of producing gold. In times of depression, the various costs of producing gold declined and it became profitable to produce gold and sell it at the fixed price. Thus more gold tended to be produced at times when an increasing quantity of it might ease financial conditions and promote recovery. Similarly, the costs of producing gold would rise in prosperity while its price would not, and gold production would be checked at the time when its continued production and use in the monetary system might permit the boom period to develop to a dangerous extent. With the value of gold relatively stable, the value of money based on gold was supposed to achieve stability as well.

Because gold could move freely from one country to another, the use of the gold standard was also supposed to equalize price levels between countries. If prices were much higher in the United States than in other countries, the other countries would purchase comparatively few things in this country, while the United States would import heavily from the other countries. Foreign exchange which could be used for making payments abroad would be at a premium in this country, while dollar exchange would be at a discount in other countries. When these premiums and discounts became so large that it would be cheaper to ship gold than to

depend upon bills of exchange (credit instruments) for making international payments, gold would flow from the United States to the other countries. The reduction of our gold stock, and the resulting contraction of money and credit in circulation, would lower the price level in this country, other things being equal, while just the opposite effects would be produced in the other countries. With this interaction the former disparity in price levels tended to be corrected.[1]

Finally, the gold standard was praised as a natural or automatic monetary system. Gold was produced or not produced as individuals desired. It could move freely into or out of monetary uses. The movements of price levels occurred slowly and gradually as a result of the myriad economic activities of millions of people, and were not contrived by governments to accomplish some economic or political objective. Changes which occurred in financial affairs, domestic or international, furnished their own corrective influences. Exchange rates were stable, and foreign trade and international investments could be carried on with a minimum of risk. The gold standard furnished, on the whole, a laissez-faire monetary system and not a managed system.

Shortcomings of the Gold Standard. The claims in behalf of the stability of value or purchasing power of gold may be admitted to this extent: the value of gold has been more stable than that of most individual commodities. But its value and that of the money based upon it have not been nearly so stable as would be desired of an article which is to serve successfully as a standard of value. In the United States the purchasing power of the dollar based on gold has varied greatly over long periods of time. For example, its purchasing power was over three times as great in 1896 as in 1920, and about twice as great at the bottom of the great depression after 1929 as in 1920. Its purchasing power has also changed significantly from one year to another. Gold production itself has not been particularly stable, and the use of gold for various nonmonetary purposes has not only varied considerably from one time to another but in a way which has been anything but closely correlated with needs for gold in monetary uses.

It is admitted that the use of the gold standard in various countries affords great stability of foreign exchange rates and that this stability is useful for purposes of international trade and investment, but this tells only half the story. Only by sacrificing stability of value within its borders can a country achieve this stability of value for its monetary unit in terms of foreign currencies. The stability of foreign exchange rates depends upon shifts in international trade, gold movements between countries, and changes in price levels within the various countries. Thus, unless a country

[1] This discussion is greatly oversimplified. (The same topic will be dealt with at greater length in Chapters XXVI and XXVII.) However, under the assumed conditions, the principles involved are valid. Of course, the gold-flow mechanism would operate in either direction between countries on the gold standard.

is willing to have its domestic price level move up and down under the influence of international conditions, it cannot maintain the gold standard in the strict sense or keep the purchasing power of its monetary unit stable in terms of foreign currencies. Critics of the gold standard often argue that, in view of the fact that for most countries international trade is only a small fraction of domestic trade, it is a poor bargain at any time to give up stability of value for the monetary unit within the country in order to stabilize its external purchasing power.

Finally, critics point out that the monetary system of a country is seldom left free to operate automatically on the basis of the gold standard. They point out that various obstacles to trade, such as tariffs, quotas, and embargoes, interfere with the international workings of the gold standard. Countries arbitrarily change the price of gold and the weight and fineness of their coins. Even under the gold standard, governments cause central banks to raise and lower interest rates, buy and sell securities in the open market, and engage in other activities which may affect, directly or indirectly, both the domestic and the external purchasing power of the monetary unit. Finally, short-term movements of capital funds between countries play havoc with a gold standard system based upon the idea that movements of funds between countries will be limited to those necessary to settle international trade balances.

The Gold Standard in the United States. Although the regular coinage of the silver dollar and the free coinage of silver in general had been abandoned many years before, it was not until 1900 that the Gold Standard Act was passed to place the United States officially on the gold standard. After 1900, the weight and fineness of the gold dollar remained the same as before, but the gold dollar was definitely established as the standard of value of the country and all forms of money were convertible into gold. Our monetary system on the basis of the gold standard was far from perfect, but in the opinion of many observers this system was superior to those which had been tried in earlier periods of our history. At any rate, we got along more or less successfully on the gold standard until 1933.

The business depression in the United States in the years following 1929 was exceedingly severe and by 1933 our government decided, as did the governments of many other countries, that recovery from the depression might occur more readily if the domestic price system were freed from the deflationary influence of world gold prices. Accordingly a great many changes were introduced into the monetary and banking system of this country in order to convert the system into one which could be managed for the attainment of certain economic objectives considered desirable by the administration and to promote recovery by raising prices. These changes resulted in the abandonment of the gold standard.

THE PRESENT MONETARY SYSTEM OF THE UNITED STATES

The Nationalization of Gold. Under the Emergency Banking Act of 1933, the Gold Reserve Act of 1934, and several interim orders and proclamations by the President of the United States, all gold coin and bullion in the United States was brought into the possession of the federal government. All private persons, companies, and institutions were required to turn their gold coin, gold bullion, and gold certificates over to the government and receive payment for them in other money. They are forbidden to own gold coin or gold certificates, and can own gold bullion only under license from the Secretary of the Treasury for purposes of export or manufacture. Even gold held by federal reserve banks and agencies was required to be handed over to the government, but the federal reserve banks were allowed to own gold certificates in the place of their former gold coin or bullion. Violation of the various provisions was made punishable by the forfeiture of the gold illegally held and by a penalty equal to twice the value of the gold. Gold coins are no longer made and old gold coins received by the government have been made into bars and are now held as bullion.

The Elimination of Gold Payments. An embargo was placed on the exportation of gold, by executive order, in April, 1933. The earmarking of gold for foreign owners and the exportation of gold coin, gold bullion, or gold certificates was forbidden until further notice, except as specifically permitted by the Secretary of the Treasury. Later, under the Gold Reserve Act of 1934, the exportation of gold in the settlement of international obligations was sanctioned once more, but only under regulations to be set up by the Secretary of the Treasury. In June, 1933, came the famous abrogation of the gold clause. A joint resolution of Congress held the gold clause, which had commonly appeared in government and corporate bonds and which provided for their payment in gold coin of standard weight and fineness, to be contrary to public policy. It was provided that henceforth all public and private obligations containing the gold clause might be paid off dollar for dollar in legal tender, which in 1933 had been made to include all coins and currencies of the United States. Finally, the Gold Reserve Act of 1934 provided definitely that the redemption of the various kinds of money in the United States in gold should be discontinued. Such redemption had in fact been discontinued under executive order since March, 1933.

The Devaluation of the Dollar. Other important monetary measures were the Thomas Amendments to the Agricultural Adjustment Act of May, 1933. These amendments empowered the President of the United States (1) to require the Federal Reserve System to buy up and hold $3,000,000,000 worth of government bonds for the dual purpose of increasing the reserves of member banks and providing a broader base for the expansion of bank credit; (2) to cause the Secretary of the Treasury to

issue up to $3,000,000,000 in United States notes (greenbacks) of full legal tender; (3) to accept from other countries on account of war debt payments an amount not to exceed $200,000,000 in silver at a price not greater than 50 cents per ounce; and (4) to fix by proclamation the weight of the gold dollar (and of the silver dollar) at any amount not below 50 per cent of its existing weight in order to stabilize domestic prices or guard the foreign trade of the United States against the detrimental influence of depreciated foreign currencies. Of these powers, the fourth is most important for our present purposes.

The president's power to devalue the dollar was further limited before he was ready to use it. The Gold Reserve Act of 1934 provided that any new weight set upon the gold dollar must fall between 50 and 60 per cent of its former weight. Immediately after the passage of this act, the president exercised his power and changed the weight of the gold dollar to 15 $\frac{5}{21}$ grains of gold nine tenths fine, which was 59.06 per cent of its former weight of 25.8 grains.[2] The avowed purposes of the devaluation were to raise the domestic price level and to stimulate the foreign trade of the United States by putting the dollar back more nearly on a par with depreciated foreign currencies. The reason for the exact extent of the devaluation is not so well known but it is strongly suspected. Since the general price level in the United States had fallen during the depression to about 59 per cent of the 1926 level, it is thought that the president hoped to restore the 1926 price level by reducing the gold content of the dollar to 59 per cent of its former weight. The reduction in the gold content of the dollar did make it possible, of course, to increase the quantity of money and credit by almost 70 per cent on the basis of a given physical amount of gold held in reserve, but the desired effect upon prices would have been forthcoming only if such an increased amount of money and credit actually got out into circulation where it could affect prices. It is generally agreed that the devaluation of the dollar actually affected the domestic price level and foreign trade of the United States only very slightly, though it did lay the foundation for a tremendous possible inflation of prices later on.

Since an ounce of gold contains 480 grains, the reduction of the pure gold content of the dollar from 23.22 grains to 13.71 grains resulted in an increase in the value or price of pure gold from $20.67 per ounce to $35.00 per ounce. This meant that the total dollar value of the gold reserves of the United States increased overnight by a little more than $2,800,000,000. While this enormous "profit" was only the result of a bookkeeping transaction, it was real money under the new definition of the dollar and could have been used for any governmental purpose. Actually, the Gold Reserve Act of 1934 had provided for the creation of a $2,000,000,000 stabilization

[2] The pure gold content of the dollar was reduced from 23.22 to 13.71 grains by this action.

fund out of the profit gained by devaluation, and this fund was set up forthwith under the control of the Secretary of the Treasury. It was to be used to deal in gold, foreign exchange, other credit instruments, and securities in order to stabilize foreign exchange rates and control the value of the dollar at home and abroad.

The Gold-purchasing Policy. The price of $35 per ounce of pure gold, which was established in the United States as the result of the devaluation of the dollar, was well above the world market value of gold and, since the government was prepared to buy all gold which was offered to it at this price, the gold stock of the United States increased by leaps and bounds. Gold production at home and abroad increased considerably, gold came out of hoards held in foreign countries, gold fled to the United States from the danger of war conditions in Europe, and countries struggling to remain on the gold standard were deprived of their reserves. Still the flood of gold into the United States continued. Our monetary gold stock, which had been a little over $4,000,000,000 at the end of 1933, increased to $12,760,000,000 in 1937 and $22,737,000,000 at the end of 1941. After 1941 our gold stock declined to some extent but it still amounted to $20,088,000,000 on August 31, 1945.[3] Gold holdings outside the United States are now relatively small and our government owns around 70 per cent of the world's gold. The gold here has been made into bars and placed respectfully in a hole in the ground which the government maintains for the purpose. With all countries off the gold standard and not likely to return to it, no one seems to know just what useful purpose the enormous gold stock of the United States will serve in the future.

The Silver-purchasing Policy. Besides all the other changes which have been noted, the United States has gone in for silver purchasing on a large scale in recent years. In 1933 the president issued a proclamation which required the mints to purchase for coinage into silver dollars any silver thereafter mined in the United States. The price was to be 64½ cents per ounce, or $1.29 per ounce with 50 per cent retained by the government as profit. Later, the Silver Purchase Act of June 19, 1934, required the Secretary of the Treasury to purchase silver both in the United States and abroad, at any time and at prices which seemed reasonable to him, within certain limits. The price to be paid per ounce could not exceed the amount of money into which an ounce of silver could be made, or $1.29 per ounce of pure silver. For silver already on hand in the United States it could not exceed 50 cents per ounce.

The Act declared the policy of the United States to be to accumulate silver until the monetary stock of that metal should amount to one fourth by value of the entire monetary stocks of the country; that is, until the

[3] Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, October, 1945, p. 1029.

value of the monetary stock of silver was one third that of the monetary stock of gold. The silver purchased could be sold again when, if ever, the market price of silver per ounce exceeded its monetary value or the monetary stock of silver exceeded in value 25 per cent of the monetary stock of both gold and silver. Silver certificates were to be issued to represent the silver acquired. They had to be issued to an amount at least equal to the cost of the silver purchased and might be issued up to its full monetary value. Under the act the president was empowered to nationalize silver at his discretion and he exercised this power in August, 1934. An embargo had been placed on the exportation of silver two months before, and the August proclamation instructed all owners of silver to turn it over to the government within ninety days at a price of 50 cents per ounce of pure silver. However, the government kept the price of silver newly mined in the United States at a much higher level. The price was 64½ cents per ounce until July 6, 1939, when the Domestic Silver Purchase Law fixed it at 71.11 cents.

Under the silver-purchasing policy, the price paid by the government for silver has been well above the world market price. Tremendous quantities of silver have been added to our monetary stock, but the announced goal of the policy has not been closely approached. Up to the middle of 1940, the government purchased some 2,370,000,000 ounces of silver to increase its stock to 3,100,000,000 ounces, but another 2,400,000,000 ounces would have been necessary to attain the stated objective.[4] While our gold stock has declined somewhat from its peak level, it still remains very large and there seems to be little chance that our monetary stock of silver will become one fourth of the value of our total monetary stocks. This conclusion should not be accepted with regret, however, for no one seems to know just what we can do with our enormous stock of silver except to hold it indefinitely. In fact, the whole silver-purchasing policy is widely regarded as a purely political expedient designed to satisfy the silver producers and the silver-producing states.

What Is Our Present Monetary System? The major changes which have been made in the monetary system of the United States in recent years have been examined, but the question remains as to what kind of monetary system this country now has. On the basis of our description of the gold standard, the United States is definitely not on this standard at the present time. The weight of our standard monetary unit has been fixed since 1943 at 15 $\frac{5}{21}$ grains of standard gold or 13.71 grains of pure gold, but the price at which the government buys and sells gold may be changed by the Secretary of the Treasury "when such action is deemed to be in the public interest." Gold is held only as bullion and is not coined. Far from

[4] R. G. Thomas, *Our Modern Banking and Monetary System*. New York: Prentice-Hall Inc., 1942, p. 734.

being full legal tender, gold cannot be used as a domestic medium of exchange. The buying and selling of gold by the government is discretionary and not obligatory. Finally, the various kinds of money in actual circulation cannot be redeemed in gold upon request, debts and other contracts calling for payments in gold cannot be enforced, and gold can be exported only as the Secretary of the Treasury permits.

Many fancy names have been coined to describe the present monetary system of the United States, but it may as well be called a managed gold reserve standard. Some reference to gold in the title is desirable, since the monetary unit of the country is defined in terms of gold, this unit serves as the standard of value of the country, and the other forms of money are backed up by large gold reserves. On the other hand, the name suggests that the gold is used only as a reserve in so far as domestic monetary affairs are concerned and may serve to give the people confidence in the kinds of money which are actually in circulation. The characterization of our monetary system as managed depends on a number of factors. It is possible for the government to vary the purchasing of silver, to issue silver certificates on purchased silver at any rate from actual cost to $1.29 per fine ounce, to devalue the silver dollar, and to manipulate the stabilization fund. Moreover, while the power of the president to vary the weight of the gold dollar lapsed in 1943, the price per ounce of gold can still be changed by the Secretary of the Treasury, as noted above.

KINDS OF MONEY IN THE UNITED STATES

We shall now examine briefly the various kinds of money which are included in the monetary system of the United States. Table 20 presents a list of these kinds of money together with the amounts in existence and in circulation. Since some of the types of money shown are held as reserves against other types, a grand total of all types of money in existence has no great significance and is not shown. The total amount of money in actual circulation, that is, money not held by the Treasury or the federal reserve banks and agents, was $27,686,000,000 as of August 31, 1945, but this total included small amounts of some kinds of money which were called in years ago and which may now be presumed to be lost or destroyed. This total of money in circulation had increased steadily since June 30, 1940, when the amount was $7,847,000,000.

Gold and Gold Certificates. All of the gold shown in Table 20 is held in the form of bullion by the United States government and none is coined or in circulation. It is supposed to include all the gold bullion in the country, save that which is held under license by private persons or firms for manufacturing or exportation. The gold certificates are representative paper money and are really just warehouse receipts for a 100 per cent reserve of

Table 20: Kinds and Quantities of Money in the United States, August 31, 1945*

Kind of Money	Amount in Existence (*in millions*)	Amount in Circulation (*in millions*)
Gold bullion	$20,088	—
Gold certificates	17,978	$ 52
Silver dollars	494	129
Silver bullion	1,588	—
Silver certificates	1,927	1,755
Subsidiary silver coins	838	808
Minor coin	309	299
United States notes	347	322
Federal reserve notes	24,462	23,685
Federal reserve bank notes	522	517
National-bank notes	120	119
TOTAL		$27,686

* Source: Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, October, 1945, p. 1029.

gold bullion. Thus the ten-dollar gold certificate (when in actual circulation) had on its face the statement, "This certifies that there have been deposited in the Treasury of the United States of America ten dollars in gold coin payable to the bearer on demand." The gold certificates were called in by the federal government in 1934 along with gold coin and bullion. They are now held by the Treasury, and by the federal reserve banks, which are permitted to keep them in place of gold in cases in which the law requires them to hold gold reserves behind paper money or credit. Some $52,000,000 of these certificates are still said to be in circulation. Some of them are still probably held in currency hoards, although heavy penalties were provided for holding these certificates in violation of governmental orders, but many have undoubtedly been mislaid or destroyed.

Silver and Silver Certificates. The silver dollars are coins containing 412.5 grains of silver nine tenths fine, or much less than a dollar's worth of silver. They are still allowed to circulate even though stocks of silver as metal have been nationalized, but only $129,000,000 of them remain in circulation, because they are regarded by most people as too inconvenient to use as a medium of exchange. The silver bullion is the silver purchased in recent years and not converted into coin. The silver certificates are representative paper money and are like the gold certificates except that the reserve behind them is entirely silver. The one-dollar bills in common use, and some of the bills of larger denominations, are silver certificates. The one-dollar silver certificate says, "This certifies that there is on deposit in the Treasury of the United States of America one dollar in silver payable to the bearer on demand." This statement refers only to enough silver to make

a silver dollar and not to a dollar's worth of silver. The silver certificates may be exchanged for silver dollars if anyone cares to do so.

Subsidiary Silver and Minor Coins. Subsidiary silver includes half dollars, quarter dollars, and dimes, all of which contain less than their face value of silver bullion. All but $30,000,000 out of $838,000,000 of these subsidiary silver coins were in circulation at last report. Minor coins are the nickel and penny which, like the silver coins, are made of materials worth less than the face value of the coins. Out of a total of $309,000,000 of minor coins, only $10,000,000 are not in circulation.

United States Notes. The United States notes or greenbacks were first issued during the Civil War to help pay the expenses of the hard-pressed federal government. They were simple promissory notes of the federal government, though they bore no interest. Their nature was clear from the words which appeared on the face of each piece, which said (on a five-dollar note), "The United States of America will pay to the bearer on demand five dollars." There was no indication that the government was holding any other money for the redemption of these notes, or of where the money was to come from. Originally, of course, there was no reserve of any kind held behind the greenbacks, and they were completely inconvertible, fiduciary paper money.

As the United States government prepared to resume specie payments in 1879, a gold reserve was accumulated for the redemption of the greenbacks. The Gold Standard Act of 1900 provided for a gold reserve of $150,000,000 to be held behind these notes. Since 1915, this gold reserve has been slightly larger and now stands at $156,039,431. The total quantity of the greenbacks has remained at $346,681,016 since 1878, when their retirement was stopped by Act of Congress, so that the gold reserve amounts to about 45 per cent of the notes in existence. The greenbacks are now inconvertible again but they circulate as freely as other types of money. However, in the interest of simplifying our monetary system, they are one kind of money which might well be withdrawn from circulation and replaced by federal reserve notes. About $25,000,000 of the available greenbacks were not in circulation as of August 31, 1945.

Federal Reserve Notes. The federal reserve notes are our most important type of paper money in the United States and make up over five sixths of all the money in actual circulation. They are direct obligations both of the federal government and of the federal reserve banks which issue them. Federal reserve notes are obtained by the federal reserve banks from the federal reserve agents [5] by depositing a like amount of gold certificates, or gold certificates and discounted or purchased eligible commercial paper (promissory notes, drafts, and acceptances), or direct obligations of

[5] A federal reserve agent is one of three governmental representatives on the board of directors of each federal reserve bank.

the United States government. The federal reserve banks themselves are required to furnish a minimum reserve of 25 per cent in gold certificates behind the federal reserve notes in actual circulation. Originally this 25 per cent reserve had to be held in gold, but when all gold was brought into the hands of the government in 1934, it was provided that gold certificates could be used in the place of gold. The federal reserve notes could be redeemed in gold at the Treasury and in gold or lawful money at any federal reserve bank until 1933. Now, of course, they can be redeemed only in other types of inconvertible and fiduciary money, and not in gold. Some $777,000,000 out of the total of $24,462,000,000 of federal reserve notes were not in circulation as of August 31, 1945. We shall deal with the federal reserve notes at somewhat greater length in the following chapters on banking.

National-Bank Notes. The national-bank notes were once the leading type of paper money in the United States and amounted at one time to a total of almost $1,000,000,000. They were issued by the national banks organized under the National Bank Act of 1863. In form they were mere promises of the respective national banks to pay lawful money on demand, but they were actually well secured. They had behind them a 100 per cent reserve in the form of government bonds of certain issues, owned by the national banks but pledged with the Treasury, plus a 5 per cent redemption fund of lawful money. However, for reasons which will be seen later, they were not a completely satisfactory type of money.

The national-bank notes are now well on their way out of circulation. In 1935 the Treasury called in for redemption all remaining government bonds which could be used as a basis for issuing national-bank notes, and these notes are being withdrawn from circulation as fast as they come in to the Treasury. However, any large issue of currency takes a long time to come in to the Treasury for redemption in other forms of money and by August 31, 1945, there were still about $120,000,000 of these notes which had not been redeemed. In all probability several million dollars' worth of these notes have been lost or destroyed and will never be heard from by the Treasury. The retirement of the national-bank notes is a move toward the simplification of our complicated monetary system and, as such, seems desirable.

Federal Reserve Bank Notes. The federal reserve bank notes are another rather unimportant item in the monetary system of the United States. They were intended originally to replace the national-bank notes. The Federal Reserve Act, passed in 1913, permitted any national bank which wished to retire all or part of its circulating notes to sell through the Treasury to some federal reserve bank the government bonds which were security for these notes. The national-bank notes would then have been retired as fast as they came in and the federal reserve banks which bought the government bonds

on which the notes were based could have issued federal reserve bank notes to replace the national-bank notes. Thus the federal reserve bank notes, when they were issued, were like the national-bank notes as to reserves, redemption, and other conditions of issue.

In practice, relatively few of the national banks desired to retire any great amount of national-bank notes, and the federal reserve bank notes amounted to very little until 1933. In that year, the Emergency Banking Act provided for the issue of federal reserve bank notes on the basis of any direct obligations of the federal government, or any eligible promissory notes, drafts, bills of exchange, or bankers' acceptances, up to the par value of the direct obligations of the government and up to 90 per cent of the face value of the eligible commercial paper. Federal reserve bank notes in actual circulation now amount to $517,000,000 out of a total of $522,000,000. They are inconvertible paper money and, like the national-bank notes, are now being redeemed in other forms of money as rapidly as they come in to the Treasury. They are now obligations of the Treasury and not of the federal reserve banks, and the issuance of new federal reserve bank notes has been stopped since June 12, 1945.

MONEY AND ECONOMIC SYSTEMS

Although we are interested primarily in the monetary system of the United States rather than in those of other countries, the principles of money which have been set forth in this chapter may be expected to apply to a great extent to other types of economic systems. By having individuals produce according to their ability and consume according to their needs by helping themselves to public stores of goods, theoretical communism, of course, aims to get along without using money or a system of prices. However, there are many people who, unless they actually see it accomplished, will not believe that an economic system can be made to operate in this way. The type of economic system which modern socialists advocate would apparently have a monetary system, and money is used in Soviet Russia and was used in the fascist economies.

In any economic system in which money is used it is likely to serve as a standard of value, medium of exchange, store of value, and basis of credit. The qualities which fit an article to function as money are not likely to change greatly from one system to another. The various general types of money which have been described might be used in any economic system. The conditions which a country must fulfill in order to be on a gold standard would not change from one economic system to another, though systems in which the government or state is all powerful in controlling economic affairs may be less likely than capitalistic systems to base their money on a particular variety of money metal. If prices are left free to rise or fall, a great expansion in the quantity of money and credit in circulation,

other things remaining the same, will tend to raise prices in any economic system. However, in Soviet Russia or in other controlled economies, with prices under governmental control, the amount of money and credit in circulation may be increased considerably with respect to the volume of trade without raising the price level. In such cases the people simply have more of the medium of exchange than they can use in purchasing available quantities of goods at controlled prices. Finally, controlled economies may seek to insulate their domestic monetary and price systems from international influences as far as possible. This may involve carrying on foreign trade as a governmental monopoly, imposing severe restrictions on imports, exports, and the use of foreign exchange, or the maintenance of separate monetary units for domestic and foreign trade.

QUESTIONS AND PROBLEMS

1. "A definition of money which served satisfactorily a few years ago may prove defective today in view of monetary conditions in the United States." Show whether you agree.

2. "Since checks, notes, and drafts are widely use in effecting exchanges, they should be classified as money," Do you agree? Explain.

3. Show how the function of money as a standard of value differs from its function as a medium of exchange.

4. "Money merely facilitates exchange transactions, and its use does not change the fact that trade is fundamentally barter." Show whether you agree.

5. Explain why the following articles would or would not make satisfactory money commodities: platinum, iron, wheat, furs.

6. "Paper money is always fiduciary in character, and metallic money may be." Show whether you agree.

7. "A creditor who is offered legal tender must either accept it or else relieve the debtor of the necessity of paying his debt." Do you agree? Explain.

8. "Paper money may be both representative and inconvertible at the same time." Explain.

9. "Paper money, though inconvertible, may be as sound and useful as any other kind of money if it is not overissued." Show whether you agree.

10. "It is desirable to hold large metallic reserves behind paper money because such reserves make it certain that the value of the money will not change." Do you agree? Explain.

11. "Other things equal, the value or purchasing power of money is a function of its quantity." Discuss.

12. Explain what is meant by the gold standard.

13. "In view of its many advantages, it is difficult to understand why the gold standard has been so widely abandoned." Discuss.

14. "Under the gold standard the monetary unit of a country is able to enjoy international stability of value only by sacrificing its stability of value within the country." Show whether you agree.

15. How has our monetary system been changed in recent years with regard to the use of gold? Explain.

16. "In 1934, the United States dollar was devalued. This act was bound to result in a sharp rise both in the price of gold and in the general price level." Show whether you agree.

17. "The gold stocks of the United States rose rapidly after the devaluation of the dollar." Why?

18. "The silver policy of the United States since 1934 has been successful in attaining the objectives which led to its adoption." Do you agree? Explain.

19. "The United States at the present time is neither completely on nor completely off the gold standard." Explain.

20. Outline the types of money which are in use in the United States at the present time.

21. "The principles of money are valid in all types of economic systems." Do you agree? Explain.

REFERENCES FOR FURTHER READING

Bradford, F. A., *Money and Banking.* New York: Longmans, Green, and Company, rev. ed., 1941, Chaps. 1–4, 11, 25–29.

Kemmerer, E. W., *Kemmerer on Money.* Philadelphia: The John C. Winston Company, 1934, Chaps. 1–12.

Laughlin, J. L., *Money, Credit, and Prices.* Chicago: University of Chicago Press, 1931, Chaps. 1, 2, 9, and 16–19.

Peterson, J. M., Cawthorne, D. R., and Lohman, P. H., *Money and Banking.* New York: The Macmillan Company, 1941, Chaps. 1–4, 20–23.

Robertson, D. H., *Money.* New York: Harcourt, Brace and Company, 1922, Chaps. 1–3.

Thomas, R. G., *Our Modern Banking and Monetary System.* New York: Prentice-Hall, Inc., 1942, Chaps. 1–3, 31–36, 46–49.

Westerfield, R. B., *Money, Credit, and Banking.* New York: The Ronald Press Company, 1938, Chaps. 1–8, 27–29, 38–40, 43–44.

XXIV

Banking

Although the study of money is important for the understanding of the way in which our economic system operates, we must recall at this point that most exchange transactions in this system are carried on today without the use of actual money. Some estimates have it that over 90 per cent (by value) of the business of the United States is now carried on by means of credit and credit instruments rather than money. Thus, while money forms the basis of credit and deferred payments, it is necessary to examine credit operations, and the banks through which these operations are carried on, in considerable detail.

THE NATURE OF CREDIT

Credit in General. On the surface, credit operations appear to be of many kinds, but they all have a fundamental similarity. In credit transactions, one party to the transaction, the creditor, turns over to the debtor a certain amount of money, commodities, or services at the present time and relies on the debtor to repay an equivalent amount (usually of money) in the future—plus an added sum, called interest, in most cases. The debtor, of course, receives money, commodities, or services at present and agrees to repay an equivalent sum of money plus interest at some future time. Thus when an individual secures a "personal" loan from a "loan shark," personal finance company, or small-loans department of a bank, he receives an amount of money which gives him immediate command over desired commodities or services and must repay the borrowed sum plus interest in a series of installments which fall due at regular intervals in the future. The purchase of an automobile, radio, electric refrigerator, or some other article on the installment plan is a similar transaction, except that the would-be debtor receives an actual commodity at present instead of a sum of money.

When the businessman borrows at a commercial bank on his own promissory note, or on the strength of a note or draft which he has received following the sale of merchandise, he receives purchasing power in the form of money or a credit to his checking account which obligates

the bank to pay money on demand. This enables him to pay wages, purchase materials and supplies, and pay taxes and other expenses in the current month or two, and at the end of the period he must make repayment. A corporation sells its bonds, and the purchasers turn over money or checks (claims against money) which will enable the firm at the present time to build an additional factory or purchase and install new machinery. Year after year the corporation pays interest to the bondholders for the use of their funds and, at the end of the stipulated period of years, it must return the principal to them intact. Even the government, in paying out fiduciary money for the commodities and services which it purchases, becomes a debtor, and the sellers of the commodities and services (and other people later on) become its creditors. That is, the government receives economic goods and in exchange for them gives out its own promises to pay instead of paying gold or some other article whose worth is equal to that of the goods purchased. However, the government's promises to pay are ordinarily so generally acceptable in exchange that they qualify as money even though they are really credit instruments in a sense. Other credit instruments which we shall discuss lack this quality of general acceptability.

Types of Credit. Credit may be classified in several ways. In the first place, there is the rather obvious distinction between consumers' credit and producers' credit. Consumers' credit involves advances of purchasing power or economic goods to consumers. Its distinguishing characteristic lies in the fact that the things acquired by the debtors as a result of the loans are not supposed to furnish them with the means of repaying the loans, but only satisfy some wants of the borrowers. The loans must be repaid out of other income of the borrowers. Under consumers' credit it is necessary to list loans made by pawnbrokers, "loan sharks," small-loans companies, and personal-loan departments of commercial banks; most of the advances made by finance companies in connection with installment selling; and, according to some people, loans made to the government for the carrying on of a war or the direct relief of unemployment.

Producers' credit, on the other hand, includes all of the advances made to individuals, firms, and governmental units which are used to facilitate the carrying on of the various branches of utility creation. It is customary to divide the loans made to producers further into subgroups called commercial, investment, and intermediate credit. Commercial credit refers to loans of relatively small amounts of purchasing power, for relatively short periods of time, and for working capital purposes. Short periods of time ordinarily mean thirty, sixty, or ninety days. Small amounts of purchasing power cannot be exactly defined, of course, but the term means that a given firm's requirements of commercial credit are usually smaller than the amounts of investment credit which it needs. The borrowed funds are used for working capital purposes if they are used to pay wages, pur-

chase raw materials and supplies, pay for light, heat, and power, or take care of other current or operating expenses. The businessman or firm is not expected to borrow funds from his bank for thirty or sixty days and use the funds to purchase new machinery or start building a wing for the factory. Such investments could hardly furnish the businessman with income for the repayment of the loan within the time period for which the advance of funds was made. Commercial loans are supposed to be self-liquidating. That is, the transactions which give rise to the need for commercial credit are supposed, within the period for which the loans are made, to furnish the borrowers with funds with which the loans may be repaid.

Investment credit refers to loans of relatively large amounts of purchasing power, for relatively long periods of time, and for fixed capital purposes. The time period involved in such loans is usually a few years at least and may run to ten, twenty, or fifty years. The funds are most often obtained through the sale of securities, and it is expected that the holders of the securities will sell them to someone else if they need to recover their funds in the meantime. The funds obtained by the borrowers are used for investment in the various types of fixed capital goods, and the repayment of the loans depends upon the productivity and earning power of these fixed capital goods over a period of years. Intermediate credit refers to loans which may run from a few months to a few years. They are usually made to finance a certain productive process, but may be made to enable a piece of capital to pay for itself through its earnings for the purchaser. Cattle loans are typical of the first variety. It takes about three years to raise cattle for the market, and the enterpriser may need loans for one, two, or three years, according to the age of the stock which he acquires. In the case of the second type of loans, a bus line may require credit for a year or two during which it can make a new bus earn enough to pay the balance which remained above the down payment.

Credit Instruments: Book Accounts. In examining briefly the credit instruments which serve as evidence of credit operations, we start with the book account, which is a simple but widely used device. As used in retail merchandising, it ordinarily requires no specific promise in writing to pay for goods purchased. An individual enters a department store or clothing store where he has a charge account, selects a suit of clothes and other items to complete a new Easter outfit, and instructs the seller to charge the merchandise to his account. Periodically (usually once a month) the store sends the individual a statement of his account, and he is supposed to pay fairly promptly. However, it is difficult to fix any definite time limits on credit of this sort. Department store statements often say that the bill must be paid at some time between the first and tenth day of the month after that in which the purchases were made, but near by in fine print may be a statement that ½ of 1 per cent interest will be added to accounts which

are more than two months overdue. Although the time of payment is indefinite and the book account furnishes no specific evidence that the supposed buyer of the merchandise has approved the debt, the book accounts or accounts receivable of an enterprise may be sold, borrowed on at a discount house, or pledged to a commercial bank as security for a commercial loan. The use of the book account is by nature confined to the sale of merchandise to consumers and producers. Where the borrower actually receives money or other purchasing power, a written promise to repay is almost always necessary.

Promissory Notes. For the purposes of the Federal Reserve System, a promissory note is defined as an "unconditional promise, in writing, signed by a maker, to pay, in the United States, at a fixed or determinable future time, a sum certain in dollars to order or to bearer." As this definition suggests, a promissory note is a stern and inflexible credit instrument. It is widely used in installment selling and in connection with all kinds of personal loans. Businessmen often give a promissory note to a commercial bank when they secure short-term loans, and it is relatively easy for them to borrow at such a bank on the basis of notes received from customers. However, the promissory note is not often required in industries and businesses from customers who are expected to pay within a month or two. The term "note" is sometimes applied to credit instruments which run for several years, but there are other and better names for such instruments.

Drafts and Acceptances. The draft is a credit instrument which does not differ greatly from the promissory note, except that the initiative in connection with the draft is taken by the creditor rather than by the debtor. In the promissory note the debtor promises to pay a certain sum of money to the creditor, but in the draft the creditor orders the debtor to pay to order or to bearer a certain sum of money at a fixed or determinable future time. After the draft is accepted by the debtor it is easy for the creditor to borrow at a commercial bank on the basis of this credit instrument. The debtor accepts the draft by writing on it the word "accepted," along with his name and the date. From this time on the draft is essentially similar to a promissory note, since the debtor has agreed to pay it when due, but it is called an "acceptance." The term "trade acceptance" is reserved for drafts which are drawn by the seller of merchandise on the buyer.

Checks. A check is really a draft, payable upon presentation and drawn on a bank. That is, a person who draws a check orders his bank to pay a certain sum of money to a third party, to himself, or to the bearer of the instrument when it is presented at the bank. Although the check is considered for many purposes as a cash item, it is really a credit instrument. In theory it is payable instantly, but actually some time usually elapses between the drawing of the check and its presentation for payment at the bank. Moreover, the acceptance of a check by a seller of merchandise or a

creditor rests upon his faith in the drawer of the instrument. The check is used in a tremendous number of business transactions, but because of the risks involved in its acceptance it is not generally acceptable as a medium of exchange and does not qualify as money under our definition.

Securities. A bond is really nothing but a long-term promissory note, for in a bond a corporation or other agency promises to pay a definite sum of money to the owner of the bond at a specified future time. Since it ordinarily is to run for a period of ten years or more, the bond is an instrument widely used in obtaining investment credit. Funds for investment credit purposes are also obtained through the sale of common and preferred stocks. The objections which will immediately be raised to the classification of stocks under the general heading of credit instruments are technically quite justified because stocks are certificates of ownership in connection with corporations and are not promises to repay specified sums of money at definite future times. However, in practice, stocks come close to being credit instruments. Stockholders are often mere capitalists rather than active owners of businesses, they give up purchasing power to the corporation which permits that organization to obtain commodities and services at the present time, and they have faith that the corporation will be so managed that, by selling the stocks to third parties, they can get their purchasing power back in the future, though not directly from the corporation itself. Even writers who would not classify stocks under credit instruments often describe the marketing of stocks by investment bankers under the heading of "investment credit operations."

The Importance of Banks in Credit Operations. Although many credit instruments may be drawn directly between individuals or companies in business, the use of credit in our economic system would probably not get very far if it were not for the intervention of banks in the various types of credit operations. Stocks and bonds, which are issued to obtain funds for long-term investment purposes, are marketed largely through institutions known as investment banks. The borrowing of funds by businessmen and firms for commercial credit purposes is almost always done at commercial banks, while promissory notes and drafts received by business enterprises in connection with the sale of economic goods are usually discounted at commercial banks instead of being held to maturity by these enterprises. Even the various types of consumers' credit and intermediate credit operations often depend indirectly, if not directly, upon banks. Our discussion of the functions and methods of banks will be limited to producers' credit operations, and we shall first deal briefly with investment banking.

INVESTMENT BANKING

The Investment Bankers. The funds that are needed in our economic system for fixed capital purposes are obtained in a variety of ways. Indi-

vidual enterprisers and partners must depend largely upon their own wealth and savings, or upon direct loans obtained from other persons with funds to invest. They may be able in some cases to secure long-term funds from commercial banks on the basis of loans secured by mortgages, or by means of short-term loans which are renewed again and again so that they have the effect of long-term loans. On the other hand, corporations, though they satisfy their own needs for investment funds to some extent by reinvesting their earnings instead of paying them out to the stockholders, depend largely upon the sale of stocks and bonds when investment funds are required.

When corporations wish to sell stocks and bonds, they usually employ the services of investment banks and bankers. In some cases corporations have attempted to sell their own securities without benefit of investment bankers and occasionally have succeeded quite well by selling for the most part to customers or employees. As a rule, however, ordinary industrial and commercial corporations are not well equipped for selling securities. They lack the salesmen, the connections, the lists of prospective customers, and other things which the investment bankers are able to provide. Commercial banks are scattered all over the country, but investment banks are found only in large cities and are concentrated to a considerable extent in New York and Chicago. They handle security issues for corporations, and for governmental units as well, and serve as intermediaries in bringing together the business and governmental institutions which desire long-term funds and the individuals and organizations which have these funds to invest. The investment bankers do not expect, if all goes well, to tie up their own funds for very long in the security issues which they sell.

Investigation. The first step in marketing an issue of corporate securities is called investigation. A representative of the corporation approaches an investment bank with a proposition for the sale of the securities. Ordinarily a well-established corporation will have rather permanent connections with a particular investment bank and will sell its securities regularly through the same institution, but a new corporation or one which desires a change will have to make a new investment banking connection. The corporate representative may have to travel from one investment bank to another before he finds a concern interested in his proposition; moreover, the investment banks do not compete actively to secure a forthcoming issue of securities.

The investment banker will carefully consider all circumstances surrounding a corporation's proposition before he decides to handle an issue of securities, and it is necessary to hit upon a price, satisfactory to both parties, at which the security issue can be transferred from the corporation to the investment bank. Once a price is agreed upon, the investment bank is likely to send a corps of accountants, engineers, or other investigators

to give the corporation's business a thorough examination, and it may even insist that it be represented on the corporation's board of directors for a time. The securities are sold outright to the investment bank, which is obligated to pay the agreed price to the corporation, whatever may happen in the sale of the securities to the investing public. That is, the investment bank must get its money back as best it can by selling the securities to individuals and institutions.

Underwriting. In spite of careful investigation of the corporation and consideration of market conditions by the investment bank, there is a considerable amount of risk connected with the purchase and sale of a large security issue by an investment bank, risk which in most cases the single investment bank will not care to bear alone. In such a case the investment bank will invite other investment banks to share the risk by forming a "syndicate agreement." There are several types of such agreements, but usually each investment house underwrites or guarantees the sale of a portion of the security issue. If the securities sell readily to the investing public at the price which the original investment house had decided to charge, the other investment houses in the syndicate have nothing to do except to receive a small margin of profit for having guaranteed the sale of the securities. However, if all or part of the securities cannot be sold at the established price, the investment banks in the syndicate have to take over the unsold securities at a somewhat lower price, and keep them until they can be sold at the desired price, or until they get tired of holding the securities.

Distribution. The investment bank makes its profit out of the margin which exists between the price paid to the corporation for the securities and the price at which they are sold to the public. The size of the margin varies from one issue of securities to another, according to the type of the securities, the nature and condition of the corporation which issues them, the condition of the market for securities, and the probable expenses of selling the issue. Various methods of selling are employed. The securities may be sold by mail to a list of regular customers, by newspaper or other advertising, by a corps of salesmen, or through ordinary banks on a commission basis. The securities are purchased in the end by financial institutions such as insurance companies, banks, and investment trusts; corporations seeking investment for surplus funds or accumulating a sinking fund; and individual investors.

In the past, investment banks usually listed on the stock exchange any issue of securities which they were trying to sell and thus "made a market" for the securities. This was done by buying and selling the securities to create the appearance of great activity for the issues and by controlling the prices of the securities on the exchange. When the securities had been actively traded in on the exchange for a time and their prices had risen to

some extent, individual investors or speculators would be attracted to purchase them. From that time on, by selling more of the securities daily or weekly than they bought back, the investment houses could gradually dispose of the securities to the public. However, the activities of investment banks in making a market for securities on the stock exchange are now controlled and limited to a great extent by the Securities and Exchange Commission, operating under the Securities Exchange Act of 1934.

The Volume of Investment Credit. Table 21 shows how the volume of investment credit in the United States, as indicated by the issue of securities, has fluctuated in recent years. The item "total security issues" includes all issues except those which are direct obligations of the United States Treasury. Since some of the securities issued merely replaced other securi-

Table 21: Securities Issued in the United States, 1927–1945*
(*In millions of dollars*)

Year	Total Security Issues	Total New Security Issues	Total New Corporate Security Issues
1927	$ 9,774	$ 7,557	$4,657
1928	9,898	8,052	5,346
1929	11,513	10,093	8,002
1930	7,619	6,912	4,483
1931	4,038	3,095	1,551
1932	1,751	1,197	325
1933	1,063	720	161
1934	2,160	1,386	178
1935	4,699	1,457	404
1936	6,214	1,972	1,192
1937	3,937	2,138	1,225
1938	4,449	2,360	873
1939	5,842	2,289	383
1940	4,803	1,951	736
1941	5,546	2,854	1,062
1942	2,114	1,075	624
1943	2,174	642	374
1944	4,153	923	627
1945	2,197	1,746	1,242

* Sources: Board of Governors of the Federal Reserve System, *Banking and Monetary Statistics*, p. 487; *Federal Reserve Bulletin*, February, 1946, p. 168.

ties which had been outstanding, the next column in the table indicates the part of total securities issued which were really new, and not refunding issues. The last column shows the extent to which the new securities issued were those of corporations. There is no indication here of the amount of investment funds furnished by the reinvestment of corporate earnings or by direct loans between individuals. Two clear tendencies are shown by the data in Table 21. In recent years the total securities issued have been made up of refunding issues to an increasing extent, and of new issues to a de-

creasing extent. Many people attribute this tendency to the governmental regulation of investment banking and of the issuance of new securities, although depressed business conditions were doubtless partly responsible for it in some years. Again, the new securities have been made up of issues of governmental units and agencies to an increasing extent and of corporate issues to a decreasing extent. If we also took into consideration the great volume of securities issued as direct obligations of the United States Treasury in recent years, we should see that new issues of corporate securities have amounted to practically nothing since 1930.

The Regulation of Investment Credit Operations. Although controlled somewhat by private regulations and state legislation, until quite recently the investment credit operations of the country were free of federal regulation for all practical purposes. After exceptionally great losses were taken by security owners in 1929 and the following years, Congress conducted an investigation of investment banking, the operation of security exchanges, and related matters, and passed legislation affecting some of these activities. The Securities Act of 1933 requires a full and fair disclosure of all material facts concerning securities which are offered for sale or sold in interstate commerce or by use of the mails. Before any new issue of securities can be sold or offered for sale, registration statements furnishing such information must be given to the Securities and Exchange Commission, a body of five members appointed by the President of the United States. Severe penalties are provided for furnishing false or misleading information and for withholding material facts.

The Securities and Exchange Commission also administers the Securities Exchange Act of 1934, which requires it to register and regulate the security exchanges of the country (with some exceptions); to register the security issues traded in on these exchanges and to require periodic reports or information from the corporations which issue these securities; to obtain and make public periodic statements concerning the holdings of officers, directors, and leading stockholders of the equity securities of their companies; to regulate transactions and manipulative practices on the security exchanges; and to regulate members of the exchanges and other persons doing business through the exchanges. The Securities and Exchange Commission has other important duties to perform, but they do not directly concern us in connection with investment banking. In our study of the principles of economics, we shall not be able to investigate more fully the acts which have just been briefly described, the reforms which they were intended to accomplish, and the results which have appeared in their operation. In any case, the problems which have usually been present in the field of investment banking and credit operations will probably be of little importance so long as the federal debt continues to grow rapidly and government bonds which are direct obligations of the federal treasury

continue to absorb practically all of the savings of the citizens of the country.

COMMERCIAL BANKING

The Functions of Commercial Banks. The functions performed by commercial banks make up quite an imposing list, but we shall concentrate on a relatively small number of them. It is clear that the commercial banks serve as central agencies for the collection of funds. Millions of persons keep in the banks any funds which they do not intend to use at the moment, and the banks turn those funds over to businessmen and firms for use in the various phases of economic activity. By furnishing a relatively safe place for the keeping of funds, the commercial banks encourage thrift and saving. Again, although commercial banks in this country are no longer allowed to issue bank notes, they do furnish the medium of exchange by means of which most business transactions are carried on—deposits subject to check. Moreover, their ability to put forth this medium of exchange is not limited to the amounts of money which they have on hand or in reserve. Since the banking laws permit the holding of fractional reserves behind bank deposits, the deposits of commercial banks may be expanded until they are several times as large as the amounts of actual cash which are held in reserve.

Finally, in creating deposits for businessmen, the commercial banks perform the important function of substituting present for future purchasing power. Suppose, for example, a businessman has a stock of goods on hand which he has produced but has not yet sold. He knows from past experience that, unless something most unusual happens, these goods will be sold within the next two months and that, when sold, they will furnish him with the purchasing power which will enable him to pay all the expenses necessary to the production of additional goods. In the meantime, since he has a large amount of money tied up in the goods on hand, he is threatened with a shortage of cash for working capital purposes and it would be most unfortunate, from the point of view of operating his enterprise efficiently, if he had to wait until his goods on hand were sold and paid for before he could continue with the production of additional goods.

In this situation the businessman may seek a loan of commercial credit from his commercial bank. He writes out a promissory note in favor of the bank in which he promises to pay the bank, say, $10,000, for value received, after sixty days. If his credit is already established at the bank, he will be able to secure a loan on the basis of this note. However, the bank may require him to pledge bonds, stocks, or other "collateral" as a guarantee that the loan will be repaid when due, since the promissory note in this case is an obligation of only one person or firm. In making the loan, the bank will ordinarily give the businessman the face value of the note less

the interest charge on that sum for sixty days. If the interest rate is 6 per cent, the charge for sixty days will be 1 per cent, or $100, and the businessman receives $9,900. He is then expected to pay the bank $10,000 in sixty days. The interest charge deducted in advance in this way is known as a "discount," and the process of lending by this method is called "discounting." When the loan is arranged, the businessman may receive the "proceeds" of $9,900 either in money or as a credit to his checking account (demand deposit). If he takes money, he will probably receive federal reserve notes, or the promises of the federal reserve banks (and the government) to pay other money on demand. If he takes a credit to his checking account, he receives the commercial bank's promise to pay money on demand by honoring his checks. Thus, in either case, in exchange for his own promise to pay money after sixty days the businessman receives bank promises to pay money on demand.

The gain for the businessman in this transaction is that the promises to pay which he receives, whether he takes money (bank notes) or a credit to his checking account, may be used immediately to pay wages, buy raw materials and supplies, and meet all the other current expenses of running his business, while the promise to pay which he gives could not be used in this way, nor could the finished goods on hand. The borrowing transaction enables him to keep his business running smoothly, regularly, and efficiently, and he thinks it quite worth while to give up $10,000 *after sixty days* in order to get $9,900 *at once* in a form widely acceptable in the business community. Of course the checks which he draws against his account in the bank are not generally acceptable in a way which would permit them to circulate as money, although they are widely acceptable, among people who know the businessman, as payment for his various business expenses.

Again, suppose that the businessman has sold a considerable quantity of merchandise but has not yet received payment for it. He has a promissory note or acceptance from his customer which will be paid in thirty days. Now the businessman is in much the same position as before. He cannot use the customer's promissory note or acceptance to pay the running expenses of his business, and these expenses must be paid if the business is to continue operating. However, he can take the commercial paper to his commercial bank and obtain a loan on it. He will endorse the promissory note or acceptance by signing his name on it, and by doing so undertakes to pay the face value of the instrument to the bank if his customer fails to do so after thirty days. The bank will ordinarily make a loan on such a note or acceptance quite readily and without requiring collateral, since the instrument has become what is called "two-name paper." That is, it is an obligation of the customer in the first place and of the businessman in the second place. From the face value of the instrument the bank deducts

the interest charge for the number of days the note or acceptance has to run, and it gives the businessman the proceeds either in bank notes or as a credit to his checking account. The businessman gets bank promises to pay money *on demand,* and gives up his customer's promise to pay money *after thirty days.* This transaction is advantageous to the businessman for the reasons given in connection with our previous illustration. Other variations of commercial credit operations could be presented, but these two examples will suffice to illustrate this important function of commercial banks.

A Commercial Bank's Balance Sheet. We shall better understand the operations carried on by commercial banks if we consider the effect of various transactions on the balance sheet, or statement of assets and liabilities, of a commercial bank. If the bank has just set up in business, the only item on the right or liability side may be capital stock; that is, the amount of money received from stockholders as the investment in the business. On the left or asset side we find cash on hand, the value of the bank building if it is owned by the bank, and the value of the bank's furniture and fixtures. If the bank is a member of the Federal Reserve System, as will be explained later, it must purchase a certain amount of the stock of the federal reserve bank of its district. This stock will appear as an asset. The sum of the various assets must equal the liability to the stockholders at this point in the bank's career.

Deposits. Soon, it is to be hoped, other items will appear in the bank's balance sheet. One such newcomer is certain to be "deposits," which will appear as a liability. From the point of view of the bank, deposits are obligations of the bank to pay money to the owners of the deposits either on demand or after a certain number of days' notice. From the point of view of the depositors, deposits represent a right to obtain money from the bank either on demand or after a certain number of days' notice. Deposits which are payable on demand are known as demand deposits or checking accounts. Those which are payable only after a certain number of days' notice are called time deposits or savings accounts. As both types of deposits obligate the bank to pay money, the deposits are a liability of the bank.

The nature of the asset that increases as deposits increase depends upon the manner in which the deposits come into being. Time deposits are set up by taking money or checks to the bank, since it would not be profitable to borrow in order to create such deposits, and they result in the first instance in an increase in the bank's cash on hand. Later they will be converted into earning assets, such as securities owned by the bank or loans and discounts due the bank. Many demand deposits are similarly set up by taking money and checks to the bank, and have about the same effect on the bank's assets and liabilities, but another large part of the bank's demand deposits will be created by the bank for businessmen who borrow

either on the strength of their own promissory notes or by discounting commercial paper received from customers in the course of business. In the latter case loans and discounts will be the asset that increases as deposits grow.

Loans and Discounts. When businessmen borrow at the bank by discounting commercial paper and take the proceeds of the loans in money, no special problem is created. In discounting a $10,000 note for sixty days at 6 per cent, the bank increases one asset (loans and discounts) by $10,000, which is the amount the borrowing businessman will have to repay later. Another asset (cash on hand) is reduced by $9,900, the amount given to the businessman. The $100 net increase in assets is offset by a $100 increase in undivided profits on the liability side of the balance sheet. These undivided profits really belong to the owners of the bank and will later be paid out as dividends to the stockholders or converted into another liability called surplus. There is in this case no increase in the deposits of the bank.

On the other hand, when businessmen borrow at the bank by discounting commercial paper and take the proceeds of the loans as credits to their checking accounts (demand deposits), the bank is placed in the peculiar position of being both creditor and debtor. It is a creditor of the businessmen to the extent of the loans which the businessmen must repay, and it is a debtor of the businessmen with respect to the newly created demand deposits. Thus the loan mentioned above in this case would increase loans and discounts by $10,000. This increase in assets would be offset by an increase of $9,900 in deposits and one of $100 in undivided profits on the liability side of the balance sheet. Since the bank will have investigated the borrowing businessman and his business very carefully to see if he is worthy of being trusted for $10,000, it is to be expected that the bank will be careful to be always ready to make good on its obligation to pay the businessman his $9,900 deposit or any part of it whenever he asks for it. The usual precaution which the bank takes is that of holding "reserves" of cash behind the deposits which it may be asked to pay at any time.

Reserves. The question of reserves, however, contains a serious problem for the commercial banker, because reserves held in the form of money do not earn anything for the owners of the bank. Every time the bank increases, by discounting commercial paper, the amount of its deposits on the basis of a given amount of cash in reserve, the earnings of the enterprise increase, but the probability that the depositors can obtain their deposits in cash, if they ask for them, is reduced. Conversely, every reduction in the amount of discounting that results in demand deposits decreases the earnings of the bank but increases the probability that the owners of deposits can obtain them in cash when they desire. If a bank had a 100 per cent reserve in money behind all its deposits, its safety for depositors would be at the maximum and its earnings at the minimum. On the other hand,

if it expanded deposits by discounting as fast as businessmen wanted to borrow, but did not bother to hold any cash reserves behind deposits, its earnings would be as great as possible but the depositors would be in a very dangerous position.

Thus the banker, torn between the desire for profits and the desire for safety, has to find some compromise between these desirable but conflicting objectives. In some countries the banker is left free to decide, on the basis of banking practice, custom, and his own conscience, how much cash he should hold behind his obligations to depositors. In other countries, as in the United States, bank reserves are regulated by law and the banker is compelled to hold definite reserves behind deposits. The nature and amount of these required reserves in this country will be investigated later in our discussion of banking. However, it may be said here that safety for depositors sometimes is greater where bankers are left to their own devices in the matter of holding reserves than where their reserves are prescribed by law.

Reserves which are required by law are usually considered as minimum reserves, and the actual amount of cash which a bank holds on hand or on deposit with other banks is well above the amount specified by law. In addition the bank depends to a considerable extent upon its so-called "secondary reserves," which are composed of items which can be readily converted into cash, such as government bonds, or other easily marketable securities, and "call loans" made to brokers and secured by stocks and bonds. In addition to reserves for the protection of depositors, a country or state may operate a system of deposit insurance in which a fund is accumulated for paying off depositors in failed banks. A country's banking system may also provide a method which permits commercial banks to borrow at some central bank or banks by pledging certain assets and thus to acquire additional cash reserves which in time of emergency may be used to pay depositors. In the United States we have both a national system of deposit insurance and a method of "rediscounting" which permits member banks in the Federal Reserve System to borrow at federal reserve banks on certain types of commercial paper which have already served as a basis for member bank loans to businessmen. These and still other methods of providing safety for depositors will merit our attention later.

Elasticity and Expansibility of Commercial Credit. Safety for depositors is a desirable objective for commercial banks, but it is not the only matter to be considered in connection with the extension of commercial credit and the creation of deposits. If the processes of production and exchange are to be facilitated by the operation of commercial banks, it is necessary that these banks be able to extend commercial credit and create deposits up to any amount genuinely needed in the economic system for industrial and commercial operations. Moreover, mere expansibility, or ability of com-

mercial credit to expand and contract in total amount, is not enough. If the objective of elasticity is to be attained, increases and decreases in the volume of commercial credit must be timed in such a way as to coincide with the expanding and contracting needs of industry and commerce for this credit.

If commercial credit (and money) could be made perfectly elastic, the ratio of the total volume of money and credit to the total volume of industrial and commercial transactions requiring the use of money and credit would remain constant, however much the economic activity of the country might expand or contract. Since the value of money (and the price level) depends upon the relationship between the total quantity of money and credit and the need for money and credit, the price level of the country would be expected to remain constant under the condition of perfect elasticity of money and commercial credit, if the velocity of circulation, or rate of turnover, of money and credit remained the same. However, as we shall see, perfect elasticity is difficult to obtain.

If money and credit were perfectly elastic, they would be performing what many persons consider to be their proper or "neutral" function in the economic system. According to this point of view, money and commercial credit are useful instruments for facilitating the making and carrying out of economic decisions which it would be in the interest of society to make and carry out even if money and commercial credit had never been invented. But money and commercial credit should never affect the content and volume of these decisions and activities. That is, money and commercial credit should not be issued in such amounts and in such ways as will cause prices to rise, profits to increase, production to expand, and the overexpansion of productive facilities to take place, so that a boom period leading inevitably to depression is created. In similar fashion, money and credit should not be prime factors making for business contraction and depression. Money and commercial credit should adapt themselves to whatever volume of economic activity is generated in the economic system. They are intended to be lubricants rather than prime movers.

The Withdrawal of Deposits. The withdrawal of individual deposits may or may not have any significant effect on the condition and the balance sheet of a commercial bank. When individuals cash checks at the bank, both the cash on hand and the deposits of the bank are reduced. However, when depositors make various payments in the community by checks, the deposits of the bank and its cash on hand are not affected if the receivers of the checks have their own accounts at the same bank. The amounts represented by such checks are merely deducted from some individual accounts and added to others. Even if the persons who receive the checks carry their own accounts at other banks in the community, the condition of the first bank is not changed significantly if there is substantial equality

between the amount of checks drawn on it and deposited in other banks in the community and the amount of checks drawn on other banks in the community and deposited in the first bank. That is, the checks drawn on other banks and deposited in the first bank will increase its deposits by just about the same amount as the checks drawn on the first bank and deposited in other banks will reduce these deposits.

Local Bank Clearings. Ingenious schemes have been devised to enable banks in the same community, and in different communities, to collect checks and other items from each other without using very much actual cash. In most cities, for example, there is a clearinghouse association to which the individual banks belong. The clearinghouse itself may be anything from a single room to a large building. Several functions may be performed by the clearinghouse, but the chief function is that of facilitating the settlement of obligations between the banks which are members of the association. For this purpose a room is maintained which contains a table, desk, or cage for the use of each member bank. The clearing, or settlement of obligations, ordinarily takes place once each day at a convenient hour, but may occur more or less often than this. As the time for clearing approaches, each bank sends two representatives to the clearinghouse. They take with them to the clearinghouse several packages of checks, one for each of the other banks. These are the checks, drawn on the other banks, which have been received by the given bank since the last clearing.

One representative of each bank is stationary and the other is roving. The stationary representatives of the various banks seat themselves at the desks or tables provided for their banks. When the gong is struck or some other signal is given for the beginning of clearing, the roving representatives pass along the row of desks or tables, leaving at each one the package containing the checks which are obligations of that bank. Each stationary representative knows in advance how much his bank is owed by all the other banks; by adding together the amounts represented by the packages of checks left at his table he soon learns how much his bank owes all the other banks. If on that day his bank owes the other banks more than the others owe his bank, the difference must be paid to the clearinghouse. If the situation is reversed, the difference is received from the clearinghouse. Since each check is an asset for one bank and a liability for another, some banks, though not the same ones from day to day, will have net balances to collect which offset the net balances which other banks have to pay. Thus the clearinghouse must break even on each clearing.

The net balances which are paid and received by individual banks may be settled with money, clearinghouse certificates (obligations of the clearinghouse to pay money on demand), checks drawn by individual banks against their balances at the federal reserve bank of the district, simple debit and credit entries to the accounts of the individual banks at

the federal reserve bank on the basis of a report from the clearinghouse as to the results of the day's clearings, or other methods too numerous to mention. In any case the economy of the clearinghouse procedure is evident. Each bank finds that its obligations to other banks almost equal the amount which it is due to receive from other banks, and the net balance which has to be settled each day is usually a small fraction of the total volume of checks cleared. The net balances to be settled between banks would probably average 10 per cent or less of the total volume of clearing transactions in most clearinghouses over any considerable period of time.

Intradistrict Clearings. Checks are used to make payments between widely separated cities as well as within a given community. For example, let us suppose that an individual in Champaign, Illinois, sends a check to a firm in Peoria in payment for some merchandise, the two cities being in the same federal reserve district. The collection of the check will be a simple matter if both the individual's bank and the Peoria firm's bank are members of the Federal Reserve System. The Peoria firm deposits the check in its bank, which is, let us say, the First National Bank of Peoria. This bank in turn relays the check to the Federal Reserve Bank of Chicago, which places a deferred credit in the account of the Peoria bank. The check is then mailed to the individual's bank (Champaign National Bank), along with a request for a remittance. This remittance may be made in various ways, or the Champaign bank may simply allow the Federal Reserve Bank of Chicago to debit its account.[1] Then the First National Bank of Peoria credits the firm's account and the Champaign National Bank, upon receiving the check, debits the account of the individual who drew the check.

Interdistrict Clearings. Now let us suppose that the member banks in the transaction are located in different federal reserve districts, which will be the case if in ordering some goods the individual in Champaign, Illinois, draws a check on the Champaign National Bank and sends it to a firm in Oakland, California. The Oakland firm deposits the check in the First National Bank of Oakland, which sends it to the Federal Reserve Bank of San Francisco. From there it is sent to the Federal Reserve Bank of Chicago, which collects it from the Champaign National Bank, and that bank in turn debits the customer's account. On the other end of the transaction, the Federal Reserve Bank of San Francisco will credit the account of the First National Bank of Oakland, which will credit the account of the firm which received the check. This explains everything except the manner in which the two federal reserve banks settle with each other.

In 1915 the Federal Reserve Board set up a fund at Washington for

[1] Both national banks in question, being members of the Federal Reserve System, will of necessity have a reserve account at the Federal Reserve Bank of the district, which may be debited or credited in clearing transactions.

the purpose of clearing items between the federal reserve banks. It was known as the gold settlement fund, and each federal reserve bank had to keep a balance of at least a million dollars in the fund. The Federal Reserve Board is now the Board of Governors of the Federal Reserve System, and the settlement fund, now known as the interdistrict settlement fund, is based on gold certificates rather than on gold, but the method of clearing remains the same. At the end of each business day each federal reserve bank wires to the Board of Governors of the Federal Reserve System a list of the credits which it has received that day for the other eleven federal reserve banks. The total amount reported by each federal reserve bank is charged to its account at the interdistrict settlement fund, while its account is credited with the amounts reported by the other eleven federal reserve banks in its favor. The total amount credited to each federal reserve bank is reported back to it, and it may compare this amount with that which it reported in favor of the other banks.

Now suppose our Champaign-Oakland transaction to be the only one on a particular day. The Federal Reserve Bank of Chicago has collected the check and has a credit to transfer to the Federal Reserve Bank of San Francisco. At the end of the day it wires the Board of Governors of the Federal Reserve System that it has a credit for the Federal Reserve Bank of San Francisco. The amount of the check is then deducted from the account of the Federal Reserve Bank of Chicago at the interdistrict settlement fund and is added to the account of the Federal Reserve Bank of San Francisco, and the item is cleared. Actually, of course, large numbers of items are cleared each day among the twelve federal reserve banks. Because commercial banks which are not members of the Federal Reserve System are permitted to participate in the schemes for intradistrict and interdistrict clearing, most items between different towns and cities are cleared in the ways just described.

THE NATIONAL BANKING SYSTEM

We now turn from the discussion of general matters in the field of commercial banking to the subject of commercial banking systems. In this connection our chief interest is in the present commercial banking system of the United States, the Federal Reserve System, but our understanding of this system will be improved if we first deal briefly with the system which preceded it—the National Banking System.

The Organization of National Banks. In 1863 Congress passed the National Bank Act which, as amended in 1864, provided the basis for the National Banking System, which endured until 1913. Under this system, five or more persons could organize a national bank, and obtain a twenty-year charter, by complying with certain requirements of the National Bank Act. The minimum amount of capital funds which had to be furnished

by the owners of a national bank varied from $50,000 to $200,000, according to the population of the town or city in which the bank was to operate. At least half the required amount of capital funds had to be paid in before the bank could start business, but the rest could be furnished in five monthly payments. The stockholders in national banks were burdened with double liability. That is, in case the bank failed, the stockholders would lose their investment and in addition could be required to pay toward the bank's debts another amount equal to the par value of their shares of stock. Each bank had to make five reports per year to the comptroller of the currency and had to stand examination by agents sent out by the comptroller.

The National-Bank Notes. Before a national bank started business it had to buy and turn over to the Treasurer of the United States registered bonds of the federal government in an amount equal to one fourth of the paid-in capital funds for banks with a capitalization of $150,000 or less and one third of the paid-in capital funds for larger banks. On the basis of these bonds as security, the comptroller of the currency would then issue national-bank notes to the national bank in an amount equal to the par value of the deposited bonds. The national-bank notes were issued in denominations from $1 to $1,000, but the amounts issued in denominations of $1 and $2 were limited so that these notes might not replace silver certificates and greenbacks in circulation.

In most respects the national-bank notes were a satisfactory form of money. All of the notes were exactly the same except for denomination and the names of the issuing banks. They had ample security behind them and they were obligations of the federal government (through the bonds held in reserve) as well as of the national banks. Besides the government bonds each national bank was required to maintain a redemption fund of gold or lawful money (money issued by the federal government), equal to 5 per cent of its outstanding national-bank notes, at the Treasury in Washington. The national-bank notes passed freely and at par as a medium of exchange in all parts of the country and greatly facilitated trade and travel. However, these notes had one great weakness—lack of elasticity—from the point of view of the money and banking system as a whole.

The national-bank notes were a boon to the federal government, since they furnished the basis for a ready market for government bonds and since bonds which could be used as reserve behind the national-bank notes did not have to bear as high a rate of interest as other government bonds. However, one may wonder, at first glance, what benefits were derived by the national banks from the issue of these national-bank notes, aside from a certain amount of prestige and advertising, since they had to maintain a reserve behind the notes of 105 per cent of their value, pay a tax on the circulation of the notes, and assume the expenses of their manufacture. That is, besides pledging $100,000 worth of bonds to get $100,000 of national

bank notes, a bank had to deposit $5,000 with the Treasury as a redemption fund. However, it must be remembered that the national banks received a moderate rate of interest on the bonds pledged with the Treasury, in addition to the current rate of interest on the lending out of the national-bank notes which were secured by the bonds. The sum of these two incomes, even after being reduced by the tax on circulation and the cost of making the notes, ordinarily exceeded by a small margin the interest which the national banks could have obtained by directly lending out the funds tied up in the government bonds and the redemption fund. Of course, if the national banks had to pay a premium above par in buying bonds to pledge with the Treasury, this small margin of gain would be reduced.

Reserves. The National Bank Act also specified that the national banks were to maintain cash reserves behind their notes and deposits. Banks in New York City, the "central reserve city," were compelled to keep reserves of 25 per cent, in the form of cash on hand, behind notes and deposits. Banks in some sixteen other places, called "reserve cities," also had to carry a 25 per cent reserve behind notes and deposits, but only half of this reserve had to be held in cash, while the remainder might be deposited with banks in New York City. Banks in other places were considered to be "country banks" and were required to keep a reserve of 15 per cent behind notes and deposits. Of this amount at least two fifths had to be kept at the banks as cash on hand, while the rest could be kept on deposit in banks in reserve cities or in the central reserve city.

In spite of the fact that the intention of these reserve provisions was undoubtedly to introduce a large element of safety into the operation of the national banks, their effects, as we shall see, were in many ways unfortunate. The practice of dividing cities and banks into three classes for purposes of reserve requirements has persisted even under the Federal Reserve System. At the present time New York and Chicago are central reserve cities, and there are some sixty-two reserve cities. For reserve purposes, banks in other places, and some of them are located in rather large communities, are still considered country banks.

DEFECTS OF THE NATIONAL BANKING SYSTEM

Minor Weaknesses. It cannot be denied that the National Banking System was a great improvement over the system or lack of system which existed before 1863, but it was defective in several respects. In the first place, the National Banking System was not really a system at all. It merely provided for the chartering by the federal government of a large number of commercial banks which remained for the most part independent and uncoordinated. No central bank or group of banks was provided to control and manage the system as a whole. The collection of checks and other obligations between banks in different cities remained awkward and

costly. Moreover, the Independent Treasury System was still operated on a specie basis, and the national banks found it inconvenient to carry on fiscal operations for the federal government. Nevertheless, the main weaknesses of the National Banking System had to do with the issuance of money and deposit credit and the manner of providing safety for depositors.

Inelasticity of the National-Bank Notes. The national-bank notes became a leading type of money under the National Banking System. They were a satisfactory kind of money in many respects, but they had the fatal defect of inelasticity. Being tied securely to a 100 per cent reserve of government bonds, the national-bank notes were incapable of expanding and contracting in total volume to meet the varying needs of industry and commerce for money. Their volume could be expanded and contracted as the total quantity of government bonds available to serve as reserves increased and decreased, but there was nothing about a period of prosperity and expanding business activity which would lead the federal government to issue more bonds and hence make possible the issue of more national-bank notes. Similarly, there was nothing in a period of business contraction and depression which would cause the federal government to pay off its debts and retire its bonds so that the total volume of national-bank notes would have to be contracted.

It is sometimes suggested that periods of prosperity and expanding business were likely to be the times when, if ever, the government would pay off part of its debts, retire government bonds, and reduce the total possible volume of national bank-notes in circulation. On the other hand, periods of depression and contracting business might well lead the government to go further in debt, issue more government bonds, and make it possible for the total volume of national-bank notes to expand. Some writers even claim that the periods in which more money was needed were likely to find government bonds selling at such a premium that there would be no incentive for national banks to issue additional national-bank notes. Any kind of money which contracted in quantity when more of it was needed in the economic system, and expanded in quantity when less of it was required, would be both inversely elastic and most unsatisfactory. It is not possible to show statistically that the national-bank notes were inversely elastic, but it is clear in any case that they were not an elastic type of money.

Commercial Credit under the National Banking System. The situation with regard to the expansion of commercial loans (and resulting demand deposits) by the national banks was somewhat different from that which existed in connection with the issue of national-bank notes. The deposits resulting from commercial loans by the banks were definitely elastic. That is, these deposits were set up by discounting commercial paper at the banks, the banks were equipped to perform this function, and every period of prosperity and expanding business, which created a need for more

demand deposits as a medium of exchange, furnished an increasing quantity of commercial paper in the industrial and business world by means of which loans resulting in demand deposits could be obtained at the banks. As a result, demand deposits, which are often called deposit currency, tended to expand and contract with business conditions and the need for a medium of exchange.

However, it would appear at first glance that demand deposits, while elastic as to timing, were not very expansible under the National Banking System. The national banks, according to their location, had to keep a 25 or 15 per cent reserve behind these demand deposits, and it would therefore seem that they could grant commercial loans resulting in demand deposits only up to the point where these deposits were four times or six and two-thirds times, respectively, the amount of cash reserves kept by the banks. Beyond this point the banks apparently could not go, since the system provided no method by which the banks could obtain more reserves in times of need, either by rediscounting commercial paper or by other means. Of course, there was nothing to keep one national bank from trying to borrow cash for reserves from other national banks, but there was almost no chance that such requests would be granted at times when all banks were busily expanding loans and deposits.

In actual practice, the expansibility of demand deposits under the National Banking System was greater than would seem possible on the basis of reserve requirements. The system of permitting country banks and banks in reserve cities to redeposit part of their reserves in banks in other cities reduced significantly the actual amount of cash which had to be held behind demand deposits, and the practice of allowing the banks to count as reserves any items due from other banks and floating around in the mails in the process of collection made it easy for the banks to maintain adequate reserves on which to expand demand deposits. In fact, one of the weaknesses of the National Banking System was found in the fact that no powers were set up in the form of a central bank or agency authorized to control the expansion of commercial loans and demand deposits in periods of prosperity. However, the inelasticity of the national-bank notes, strangely enough, set some limits on the expansion of demand deposits. In periods of prosperity, when large amounts of money were needed in actual circulation and the national-bank notes could not be expanded adequately, money for use in exchange tended to be drawn out of the reserves of the banks, and this fact curtailed to some extent the possible expansion of loans which resulted in demand deposits.

Safety for Depositors. A final and glaring weakness of the National Banking System was that it did not furnish an adequate degree of safety for the depositors in national banks. The reserves required by law behind deposits were fairly large. Indeed, they were larger than those which have

usually prevailed under the Federal Reserve System. However, because of the redepositing or pyramiding of reserves, the amounts actually held in cash in the system were much smaller than the sums apparently required by law. For example, consider the case of a country bank with $1,000,000 in deposits. Behind these deposits the law required a reserve of 15 per cent, or $150,000, but only two fifths of this sum, or $60,000, had to be retained by the bank as cash on hand. The remainder could be carried with a bank in a reserve city, but the reserve city bank did not have to hold the $90,000 in cash. This sum was treated like any deposit and could be invested or loaned out, subject only to the requirement that a 25 per cent reserve, or $22,500, be kept behind the $90,000 deposit.

The reserve city bank did not have to keep even this $22,500 in cash. Half of it was held in the form of cash and the rest could be carried with a bank in a central reserve city. This $11,250 would be treated as an ordinary deposit by the bank in the central reserve city and only a 25 per cent reserve in cash, or $2,812.50, had to be held behind it. Thus the total cash reserves carried behind deposits of $1,000,000 in a country bank might be only $60,000, plus $11,250, plus $2,812.50, or a total of $74,062.50. This amounted to less than a 7.5 per cent cash reserve behind the deposits in question, instead of a 15 per cent reserve. Moreover, the $60,000 in cash which the country bank was supposed to hold might be reduced by counting as reserve "float," or items due from other banks and in the process of collection.

The result of the whole situation would be that, when the country bank needed its reserves, it did not have them and could not get them. It would, of course, ask the reserve city bank to return its $90,000, but the reserve city bank would not be holding this amount in cash and it would have to call in loans, sell investments, and so on, in order to get the cash. Incidentally, it would call on the central reserve city bank to return its deposit of $11,250, but the central reserve city bank could not comply with such requests except by calling in loans, selling securities, and so on. Thus when large numbers of banks were trying to recover their reserves, it was not surprising to find a panic in the securities markets, extremely restricted credit because of the calling in of loans, failures of business houses whose credit at the banks was cut off, and the failure of banks which could not meet the demands of depositors who asked for their deposits in cash.

The weakness of the National Banking System with respect to safety for depositors was not due entirely to the pyramiding of reserves. Even if no redeposit of reserves had been permitted and all national banks had held the full legal reserves in cash, the depositors would still have been in grave danger in times of emergency. It is clearly impossible to pay off 100 per cent of deposits with 25 or 15 per cent in cash, even if the banks could use all their legal reserves for this purpose, and the National Banking

System provided no method by means of which the national banks could obtain additional cash reserves when they were needed to meet the demands of depositors. Thus when large numbers of depositors lined up outside a bank's door to ask for their deposits in cash because they were afraid the bank could not pay them, the bank might pay some of them by using its cash, selling securities, and calling loans, but before long it would have to pull its shades, bank its fires, and close up shop, and a new name would be added to the list of bank failures.

QUESTIONS AND PROBLEMS

1. "Credit operations appear to be of many kinds, but they all have a fundamental similarity." Show whether you agree.

2. Distinguish between investment and commercial credit.

3. "A book account is not a credit instrument, strictly speaking, but it often takes the place of a credit instrument." Do you agree? Explain.

4. "A draft or acceptance is really a promissory note except that the creditor rather than the debtor takes the intiative in using the former instrument." Show whether you agree.

5. "Since a check is payable at sight, it should not be classed as a credit instrument." Show whether you agree.

6. What reason is there for considering stocks as well as bonds as credit instruments?

7. "The function of investment banks is carried out in three fairly distinct steps." Explain.

8. "Two trends are readily noted in data showing securities issued in the United States in recent years." Explain.

9. "The chief function of commercial banks is to serve as central agencies for the collection of capital funds." Do you agree? Explain.

10. Explain and illustrate the main function of commercial banks.

11. "A businessman may find it easier to obtain commercial credit on the basis of his customer's note than on his own." Why?

12. The following items appear on the balance sheet of a commercial bank: deposits, loans and discounts, surplus, balance with a federal reserve bank.

    a. Explain why each item is an asset or a liability.

    b. Explain how the deposits may have been created.

13. "In creating a deposit for a businessman, a commercial bank does not use any of its money and therefore should not be allowed to charge interest on the transaction." Show whether you agree.

14. "The commercial banker, if left to his own devices, would be in a dilemma in trying to decide how great his reserves behind deposits should be." Do you agree? Explain.

15. Distinguish between primary and secondary reserves.

16. How does the elasticity of credit differ from its expansibility? Explain.

17. "If bank credit can be expanded ten times on the basis of gold reserves, it is obviously elastic." Show whether you agree.

18. "If bank credit (and money) were perfectly elastic and expansible, the general price level would never change." Do you agree? Explain.

19. "The withdrawal of individual deposits may or may not cause commercial banks to pay out cash." Explain.

20. Describe the manner in which checks and other items are cleared between banks within a given city.

21. "It is simple to clear checks and other items between member banks which are located in different cities within the same federal reserve district." Discuss.

22. How does clearing take place between member banks which are located in different federal reserve districts? Explain.

23. "The national-bank notes furnished a sound but inelastic currency." Explain.

24. "Reserves against deposits were relatively large under the National Banking System, but they failed lamentably to provide safety for depositors." Discuss.

25. "The National Banking System was not really a system at all." Show whether you agree.

26. "Commercial credit was elastic under the National Banking System but not sufficiently expansible." Do you agree? Explain.

27. Show what was meant by the "pyramiding of reserves" under the National Banking System.

See References for Further Reading at the end of Chapter XXV.

XXV

Banking (*Continued*)

As a result of the demonstrated weaknesses of the National Banking System, a movement to obtain banking reforms was under way during the last years of its operation. Finally, in 1913, the National Banking System was replaced by the Federal Reserve System; or perhaps it would be more accurate to say that the Federal Reserve System was superimposed on the National Banking System, since the national banks continued to exist and function under the new system. The Federal Reserve System was intended in general to introduce a greater degree of centralization and unification than formerly into the commercial banking system of the United States and specifically to correct the defects which had existed in the National Banking System. Before we study the ways in which these weaknesses were to be remedied, we must examine the structure and organization of the Federal Reserve System.

THE STRUCTURE OF THE FEDERAL RESERVE SYSTEM

The Federal Reserve Districts. The whole United States is divided into twelve federal reserve districts, with one federal reserve bank in each district. These banks are located in Boston, New York City, Philadelphia, Cleveland, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas, and San Francisco. In addition, each federal reserve bank is allowed to establish branches in its district as necessary. Such branches exist in most districts and in some of the larger districts there are several branches. In setting up the federal reserve districts, attention was paid to such matters as area, population, principal types of economic activity, facilities for transportation and communication, banking needs, and the ability of the prospective member banks to contribute the necessary amount of capital funds to the federal reserve banks. It was intended that the districts should be areas of equal banking needs, however much they might differ in other respects.

The Member Banks. When the Federal Reserve System was set up and all national banks were required to become members of the new system

or else give up their charters as national banks, practically all of them decided to become members. State banks and trust companies were permitted to become members of the system by meeting the various requirements for membership. In the first place, a member bank has to have a minimum of $50,000 of capital funds contributed by its owners.[1] Second, each member bank may be required to purchase capital stock in the federal reserve bank of its district to an amount equal to 6 per cent of the member bank's capital and surplus. Dividends received by member banks on their federal reserve bank stock are limited to 6 per cent annually. Third, all member banks must keep their legal reserves behind demand and time deposits in the federal reserve banks of their districts, and these reserves earn no interest whatever. Under the Federal Reserve Act as amended, member banks in central reserve cities must keep reserves of 13 per cent behind demand deposits and 3 per cent behind time deposits; member banks in reserve cities must keep reserves of 10 and 3 per cent behind demand and time deposits, respectively; and country member banks must keep reserves of 7 and 3 per cent behind demand and time deposits, respectively. These are the basic reserves, but the Board of Governors of the Federal Reserve System may change these requirements from time to time to an extent which will be described later on.

Finally, in addition to these major requirements, member banks must abide by various restrictions and limitations imposed by the system upon loans, investments, and banking practices generally. Without going into great detail, it may be said that member banks are limited as to the percentage of their capital and surplus which may be placed in real-estate loans, in loans based upon stocks and bonds as collateral, and in loans to any one individual or firm; as to the percentage of the current market value of real estate and securities which they may lend; as to loans to bank officers or on the basis of the bank's own stocks; as to the charging of fees for the collection of checks and other items on distant banks; and as to dealing in securities.

While practically all national banks joined the Federal Reserve System at once, many state banks and trust companies have remained outside the system. As of June 30, 1945, only 6,840 banks out of a total of 14,542 were members of the Federal Reserve System.[2] However, although the banks in the system amounted only to about 47 per cent of the total number of banks, these member banks were handling about 74 per cent of the loans of all banks, about 78 per cent of the investments of all banks, and about 78 per cent of the deposits of all banks. Hence the membership figures alone do not give an accurate picture of the strength and importance of the

[1] Some state banks in communities of less than 3,000 population are required to have capital of only $25,000.

[2] Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, October, 1945, p. 1032.

Federal Reserve System. Of the member banks in the Federal Reserve System on June 30, 1945, some 5,015 were national banks and only 1,825 were state-chartered member banks. This means that almost 80 per cent of all state-chartered commercial banks were still outside the system.

Why have so many state banks remained outside the Federal Reserve System? The reasons are many and varied, but the most important ones are directly connected with the requirements for membership. Many state banks are small and have been unable to meet the minimum capital requirements for membership. In some cases they have been reluctant to invest the required amounts of funds in federal reserve bank stock because they have thought that they could make these funds earn more than 6 per cent annually in other uses. State laws have permitted the state banks to keep part of their reserves in other banks where (until 1933) they could command at least a small amount of interest, while the reserves of member banks have had to be kept in the federal reserve banks without interest. The restrictions placed on the loans, investments, and banking practices of member banks under the Federal Reserve System have been more severe than those imposed on state banks by the laws of some states. For example, many state banks derive quite a little income from exchange charges on checks presented through the mails. As members of the Federal Reserve System they would have to give up this practice and participate in the collection of checks at par. Finally, state banks have remained outside the Federal Reserve System for the same reason that some persons will not buy a bottle of ink as long as they can fill their fountain pens at the post office. That is, nonmember banks have been able to secure some of the advantages of belonging to the system without incurring the expenses and obligations which membership entails.

The Federal Reserve Banks. Each federal reserve bank was established with a minimum capitalization of $4,000,000 made up of $100 shares. This federal reserve bank stock is owned by the member banks in each district, since each member bank may be required to subscribe to the stock of the federal reserve bank of its district to an amount equal to 6 per cent of the member bank's capital and surplus. In practice, the member banks have been required to invest only 3 per cent of their capital and surplus in federal reserve bank stock, but it is nevertheless correct to say that the federal reserve banks are owned by the member banks, and are not privately owned or governmental institutions. Dividends of 6 per cent annually are paid on the stock, and any surplus earnings of the federal reserve banks must be devoted to other specified purposes.

Each federal reserve bank is controlled by a board of nine directors. Three directors, known as Class A members, represent the member banks of the district, while three others, called Class B members, represent the commercial, agricultural, and industrial interests of the district. The direc-

tors of these two classes are elected by the member banks of the district by means of a rather complicated election system. The other directors are Class C, or governmental representatives, and are selected by the Board of Governors of the Federal Reserve System. One of the Class C members is chairman of the board of directors and serves also as federal reserve agent. In the latter capacity he has charge of the issue of federal reserve notes to the bank. All directors hold office for three years and they appoint the president, vice-president, and other officers of the bank.

The federal reserve banks perform a considerable number of functions. They issue federal reserve notes; receive deposits from member banks and hold member bank reserves; carry on clearing and collection operations for member banks; rediscount eligible commercial paper for member banks and make direct loans to these banks; buy and sell bonds of the federal government, and certain other securities, in the open market; buy and sell bills of exchange which arise from commercial transactions; and set rediscount rates, subject to review and determination by the Board of Governors of the Federal Reserve System. In general, as most of these functions suggest, *the federal reserve banks are intended to be bankers' banks, and to perform those functions for the member banks which the member banks carry on for their own customers.*

The Board of Governors of the Federal Reserve System. The agency at the head of the system is called the Board of Governors of the Federal Reserve System. It consists of seven members, appointed by the President of the United States with Senate approval. The term of office is fourteen years and the salary is $15,000 per year. The members' terms expire in rotation, one in each two-year period, and members who have served a full term may not be reappointed to the board. The chairman and vice-chairman of the board are designated by the president from among the seven members and serve in these capacities for a period of four years. As some of these specifications suggest, the Board of Governors is intended to be as independent as possible of political, business, industrial, agricultural, or other interests so that it may exercise its important powers without fear or favor.

Some of the powers of the Board of Governors are routine and supervisory in character, but others are very important in connection with the objectives of safety for depositors and elasticity of currency and credit. The Board of Governors controls the open-market committee which determines policy in connection with the open-market operations of the federal reserve banks; it controls the process of rediscounting, determines what is eligible commercial paper, and for all practical purposes determines the rediscount rates charged by the federal reserve banks to member banks; it controls the amount of credit which may be granted by banks, brokers, and others on the basis of securities; it may deny member banks the use of the

rediscounting and other facilities of the system if these banks have been using bank credit for improper purposes; it controls clearing operations and the charges which banks may make for collecting checks or other items from other banks; it limits interest rates which member banks may pay on demand and time deposits; it may suspend reserve requirements for limited periods of time; and it may change the amounts of reserves which member banks are required to keep behind deposits. The significance of most of these powers will become apparent as we proceed.

ELASTICITY OF CURRENCY AND CREDIT

The Federal Reserve Notes. Under the Federal Reserve System, the federal reserve notes have become the leading type of money in circulation. As we noted in the preceding chapter, the national-bank notes had been an inelastic type of money because they were based on a 100 per cent reserve of government bonds, and these bonds did not vary in amount as the needs of the economic system for money in circulation changed from time to time. The federal reserve notes, on the other hand, were made an elastic type of money by the provision that the principal reserve behind them should consist of commercial paper, a kind of reserve which could be expected to fluctuate in total amount directly with business and industrial conditions. In explaining the process by which the federal reserve notes are supposed to move in and out of circulation, we shall assume that the proceeds of loans granted by the member banks are taken by the borrowers in money rather than as credits to checking accounts.

Since a considerable part of the exchange transactions in the United States is carried on by means of credit instruments, every period of prosperity and expanding business gives rise to an increase in the amount of commercial paper (promissory notes, drafts, and acceptances) in existence in the economic system. The business and industrial firms which receive these pieces of commercial paper seldom desire to hold them to maturity, for reasons which we examined in the preceding chapter; hence the commercial paper is usually taken to commercial banks to be discounted. According to our present assumption the borrowing firms take the proceeds of the loans in money (federal reserve notes) rather than as credits to their checking accounts. Since loans of this type create no deposits and require no reserves to be kept by the commercial banks, they may be made as long as the banks have money available for the purpose. When the member banks run short of money, they can look over the commercial paper received from borrowers and take to the federal reserve banks of their districts any part of it which is eligible for rediscounting. A rediscount is merely a second discount (by federal reserve banks) of commercial paper which has already been discounted for businessmen by member banks. From the face value of the paper the federal reserve banks deduct the

rediscount rate of interest for the length of time which the pieces of commercial paper still have to run. The member banks may then receive the proceeds of the rediscounting either in federal reserve notes or as credits to their reserve accounts at the federal reserve banks.

We shall assume at present that the member banks ask for the proceeds of rediscounting in the form of federal reserve notes. In this case the federal reserve banks turn over to the federal reserve agents the rediscounted commercial paper, plus 25 per cent in gold (now gold certificates) on the amount of federal reserve notes to be issued, and they receive a new issue of federal reserve notes which are passed on to the borrowing member banks. These member banks can then resume the lending of cash to business firms, run short of cash again, take more eligible commercial paper to the federal reserve banks for rediscounting, secure more federal reserve notes, and so on and on. Thus the federal reserve notes in circulation may increase in total amount as the needs of industry and business for money increase, and elasticity may be secured.

The limits on the process are found in the gold reserve requirement and in the eligibility requirements for commercial paper which is to be rediscounted. However, because of our large gold reserves, the gold (gold certificate) reserve requirement has not usually set any very serious limit on the expansion of federal reserve notes. In case of emergency the federal reserve notes may be issued with less than the prescribed amount of gold reserve, but the federal reserve banks must pay a tax on the deficiency of reserves. This tax becomes larger as the deficiency becomes greater. Thus in so far as the gold reserve requirement is concerned, there is no doubt of the ability of these notes to expand in amount to any necessary extent.

The eligibility requirement for commercial paper which is to be rediscounted is another matter. To be eligible for rediscount (before 1937) the commercial paper had to be paper endorsed by member banks, with a maturity of not more than ninety days at the time of rediscount;[3] and it also had to be a direct result of actual industrial, commercial, and agricultural transactions in the purchase and sale of goods. Commercial paper resulting from real-estate loans or drawn for the purpose of trading in, carrying, or purchasing securities other than government bonds was ineligible for rediscount. Thus eligibility was limited to short-term, self-liquidating commercial paper. Since 1937 the eligibility provisions have been broadened slightly. Otherwise satisfactory commercial paper may now be rediscounted if the proceeds of the loans were originally used to finance some other borrower in an actual commercial, industrial, or agricultural transaction. This extends eligibility to some commercial paper of commission merchants and finance companies. It is also now possible to rediscount short-term notes resulting from loans made to finance the construction of

[3] Except for certain types of commercial paper derived from agricultural operations.

houses and farm buildings, provided a person approved by the federal reserve bank concerned agrees to pay the full amount of the loan upon completion of the building.

We have seen how federal reserve notes come into circulation in large quantities when business is prosperous and expanding, but we must now look at the other side of the picture. When economic activity is on the down grade and commercial and industrial transactions are decreasing in volume, the amount of commercial paper in use in the economic system also declines. Businessmen pay off old loans at the banks faster than they take out new ones, and funds accumulate at the banks. As soon as the banks are able to do so, they pay off their indebtedness at the federal reserve banks and recover the commercial paper which they have pledged for loans. Thus federal reserve notes move back to the federal reserve banks when they are no longer needed in circulation.

Demand Deposits. The demand deposits of member banks were also made more elastic under the Federal Reserve System. In illustrating the process which is involved, we now assume that the proceeds of loans granted by member banks are taken by borrowers in the form of demand deposits, or credits to checking accounts at the banks. As before, we start with a period of prosperity and expanding business, an increase in the volume of commercial paper in use in the economic system, and the tendency on the part of business and industrial firms to discount the commercial paper at the banks. Since the borrowing firms now receive the proceeds of the loans as credits to their checking accounts, demand deposits are created behind which the member banks must keep reserves at the federal reserve banks of (normally) 13, 10, or 7 per cent, depending on the location of the member banks. Such loans, and the demand deposits which result, can be increased until the member banks' reserves at the federal reserve banks are reduced to the minimum required percentages of demand deposits, and then these reserves must be increased if further lending is to occur.

The member bank reserves may be increased in several ways, but the basic method is by rediscounting. That is, member banks in need of additional reserves select eligible commercial paper from their portfolios and take it to the federal reserve banks for rediscounting. In the case now under discussion it would do the member banks no good to take the proceeds of the rediscounting in the form of federal reserve notes, for this action would only increase their ability to make cash loans and not loans which result in demand deposits. The member banks therefore leave the proceeds of rediscounting on deposit at the federal reserve banks; whereupon these amounts become part of the reserves of the member banks. The member banks thus acquire the legal right to increase their demand deposits, by lending, to an amount which is 7.7, 10, or 14.3 times the reserves

obtained by rediscounting, according to whether the banks must keep 13, 10, or 7 per cent reserves behind their demand deposits.

This process, too, may be repeated over and over again. The banks may make loans and create deposits until their reserves drop to the minimum again, rediscount more commercial paper, obtain still more reserves, and so on. The expansion of demand deposits by this process is limited by the eligibility requirement for the commercial paper which is to be rediscounted and by the requirement that the federal reserve banks themselves must hold a 25 per cent reserve in gold (gold certificates) behind the reserves of the member banks. That is, the federal reserve banks can create reserves for the member banks by rediscounting only as long as the federal reserve banks can furnish the necessary reserve behind member bank reserves. The whole process is also limited, or may be, by the powers entrusted to the Board of Governors of the Federal Reserve System for controlling the expansion of credit by the member banks. In times of depression and declining business activity the demand deposits of member banks, as created by lending, are contracted as businessmen pay off their loans at the banks, and the banks get out of debt to the federal reserve banks.

When a member bank obtains reserves by rediscounting commercial paper at a federal reserve bank, it acquires the legal right to increase its demand deposits by an amount which is 7.7, 10, or 14.3 times the amount of the new reserves, depending on whether the member bank is subject to a 13, 10, or 7 per cent reserve requirement for demand deposits. This does not mean, however, that an individual member bank will always actually be able to increase demand deposits to such an extent. If just one member bank expanded demand deposits up to, say, ten times the amount of new reserves obtained by rediscounting, the business firms which received the deposits would draw a large volume of checks against them. As a result the checks which the one member bank had to pay through the clearing-house and otherwise would be much greater in total amount than the checks which it received to collect from other banks, and the one member bank would be swamped with demands for cash. When member banks in general are busily expanding demand deposits, the checks drawn on any one bank are largely canceled or offset as usual by the checks drawn on other banks which it receives. Consequently, the member banks as a group may expand demand deposits up to several times any amounts of reserves which are obtained by rediscounting.

Expansibility of Federal Reserve Notes and Demand Deposits. Closely related to the question of the elasticity of federal reserve notes and demand deposits is the matter of expansibility, or the extent to which federal reserve notes and demand deposits can be expanded on the basis of the standard money, gold. Since federal reserve notes require at least a 25 per cent reserve of gold (gold certificates), it is clear that no more than $4.00 worth

of these notes can be issued on the basis of $1.00 in gold held by the government. However, demand deposits in member banks can be expanded to a much greater extent than this. Federal reserve banks must keep a 25 per cent reserve behind the reserves of member banks. This means that $1.00 in gold held by the government, represented by $1.00 in gold certificates held by a federal reserve bank, will support member bank reserves to the amount of $4.00. Then, on the basis of this reserve, a member bank may create deposits for businessmen up to the amount of $30.77, $40.00, or $57.14, according to whether the member bank has to keep a 13, 10, or 7 per cent reserve, respectively, behind its demand deposits. Thus we see that demand deposits in member banks have an expansibility of from roughly 31 to 57 times on the basis of $1.00 in standard money, gold. Federal reserve notes and demand deposits in member banks are both quite elastic, but they have different degrees of expansibility on the basis of gold.

THE CONTROL OF COMMERCIAL CREDIT

The Expansion and Contraction of Money and Credit. In the first twenty years of the operation of the Federal Reserve System (1913–1933), it became clear that federal reserve notes and the demand deposits of member banks were readily able to expand and contract in volume as the needs of industry and trade for money and commercial credit increased and decreased. However, it also became evident that the expansion and contraction of federal reserve notes and demand deposits were not subject to adequate controls. The result was that, in periods of prosperity and thriving business, federal reserve notes and demand deposits would not merely expand but overexpand, while in periods of business depression overcontraction would occur. From the point of view of obtaining a perfect adaptation of the total quantity of money and commercial credit to the needs of business, overexpansion may be as bad as, or even worse than, underexpansion. And the same thing is true of overcontraction.

In the period from 1915 to 1920, the federal reserve notes in circulation increased from the modest sum of $188,817,000 to $3,336,281,000, the commercial paper rediscounted by the federal reserve banks increased from $32,300,000 to $2,687,393,000, and the deposits of all member banks increased from $10,636,000,000 to $24,220,000,000.[4] That these expansions of money and deposit credit were much greater than necessary to accommodate the volume of exchanges which occurred during the period is indicated by the fact that the general level of wholesale prices in the United States, on the basis of 1926 prices as 100, increased from 69.5 in 1915 to 154.4 in 1920.[5] However, the period in question was that of World War I, and it is often concluded

[4] Department of Commerce, Bureau of the Census, *Statistical Abstract of the United States*, 1939, pp. 240–250.

[5] *Ibid.*, p. 316.

that the Federal Reserve System did not function too badly under the abnormal conditions of the times.

After a short but sharp depression in late 1920 and in 1921, another great boom period occurred under the auspices of the Federal Reserve System, but this time there was no great war to serve as an excuse. The federal reserve notes in circulation had declined sharply from the 1920 peak and, in the period from 1922 to 1929, never increased again to the 1920 level. The same thing was true of the commercial paper rediscounted by the federal reserve banks. However, the loans of all member banks, which had amounted to $19,555,000,000 in 1920, increased to $26,150,000,000 by 1929; the investments of these banks grew from $5,976,000,000 in 1920 to $10,529,000,000 in 1928; and the deposits of these banks expanded from $24,220,000,000 to $39,067,000,000.[6] The wholesale price level did not change greatly during the boom period from 1922 to 1929, but the general price level, including security prices, increased considerably. By 1933, after four years of depression, the loans of all member banks had declined to $12,833,000,000, their deposits to $27,167,000,000,[6] and the wholesale price level to 65.9 on the basis of 1926 as 100.[7] These statistics strongly indicate that the Federal Reserve System furnished us with a highly flexible system of money and deposit credit, but that it lacked adequate powers for controlling expansion and contraction.

Rediscount Rate Changes. The founders of the Federal Reserve System, well aware of the fact that control over the expansion and contraction of money and credit was essential, set up powers in the system that were supposed to furnish such control. One power, to be exercised by the individual federal reserve banks, was that of raising and lowering rediscount rates in the various districts of the system. When an overexpansion of money and credit seemed imminent, the federal reserve banks were supposed to raise the rediscount rates charged on loans made to member banks. The member banks, finding it more expensive than formerly to borrow at the federal reserve banks, were then supposed to charge business and industrial firms higher rates of interest which would discourage borrowing from the member banks. In this way the expansion of money and credit was to be checked.

In practice this method of control was quite ineffective. For one thing, the rediscount rates charged by the federal reserve banks were regularly kept lower than the rates which member banks charged in lending to their customers, so that member banks continued to receive a small margin of interest even on rediscounted commercial paper. More important, however, was the fact that member banks, by rediscounting commercial paper and leaving the proceeds as reserves at the federal reserve banks, became

[6] *Ibid.*, p. 248.
[7] *Ibid.*, p. 316.

able to create deposits, by lending, to an amount greater than that which they borrowed at the federal reserve banks. An individual member bank, even if others were not expanding credit as rapidly, could create deposits which were somewhat larger than the sums obtained by rediscounting at the federal reserve bank. This was possible because the businessmen who obtained demand deposits by borrowing at the member bank would probably not draw out all of these deposits during the life of the loans, and many of the checks which they did draw would be redeposited in the same bank and require no cash outlay by the bank. If all or most of the member banks were expanding credit rapidly, they could usually create deposits which were several times as large as the reserves obtained by rediscounting at the federal reserve banks.

The significance of these facts is easy to see. If a member bank, by borrowing $1,000 of reserves through rediscounting at the federal reserve bank at 5 per cent interest, becomes able to lend $5,000 in deposits at 6 per cent interest, it will not worry much if the rediscount rate is raised from 5 per cent to 6 per cent. It will be very profitable to continue lending without raising the interest rate charged to businessmen at all, and, if the bank does decide to pass the increased cost of rediscounting on to borrowers, it can do so by raising its interest charge only one fifth of 1 per cent. On the same basis of credit expansion, it would require only a 1 per cent increase in the interest rate charged by the member bank to pass on to borrowers the cost of a 5 per cent increase in the rediscount rate charged by the federal reserve bank. Moreover, if borrowers at member banks expect to be able to make highly profitable use of borrowed funds, an increase of 1 or 2 per cent in the rate which they have to pay for these funds at the member banks is not likely to discourage borrowing to any great extent.

There are also other factors to consider. If member banks have large excess reserves when a period of credit expansion gets under way, the business boom may almost reach runaway proportions before member banks are forced to do any rediscounting, and rediscount rate changes cannot be effective unless member banks borrow at the federal reserve banks. Again, when rediscount rates are high, international conditions may lead individuals and firms in other countries to send gold to the United States for safekeeping. Member banks receive the gold, deposit it in the federal reserve banks, and the reserves of member banks are increased without rediscounting. Thus one difficulty in controlling credit expansion by changing rediscount rates is to make sure that member banks will come in to the federal reserve banks to borrow.

Open-Market Operations. This difficulty was supposed to be handled by the second major power set up for controlling the expansion of member bank credit—the power to engage in open-market operations. When an overexpansion of bank credit was threatened, the federal reserve banks

would sell, say, a billion dollars' worth of government bonds in the open market. For the sake of simplicity, let us assume that the business firms, individuals, and institutions purchasing these bonds would pay for them with checks drawn on the various member banks. The federal reserve banks could collect these checks from the member banks by charging the member banks' reserve accounts. Since the member banks have to keep a reserve of only 13, 10, or 7 per cent behind demand deposits, a reduction of one billion dollars in their reserve accounts would mean a reduction of 7.7, 10, or 14.3 times as much in the volume of demand deposits which these reserve accounts could support. Of course, the member banks would be neither willing nor able to reduce demand deposits quickly by any such amounts, so that they would set about replenishing their reserve accounts, presumably by rediscounting. Then, when they attempted to rediscount commercial paper at the federal reserve banks, they would be confronted with the high rediscount rates which the federal reserve banks would be charging at such a time. However, even if open-market operations were successful in bringing the member banks in to the federal reserve banks to rediscount, this would not necessarily make changes in the rediscount rates effective in controlling the expansion of credit, for the reasons which we have explained.

Federal Reserve Credit Policy. These two powers were all that were made available in the Federal Reserve System for controlling the expansion of bank credit, unless we count the moral suasion and advice which the Board of Governors (then called the Federal Reserve Board) and reserve banks could exercise, as well as their attempts to foster the traditional disinclination of American commercial banks toward being in debt to other banks. The member banks which conformed to the recommendations of the central banking authorities at such times would, of course, hold their rediscountings to as low a figure as possible. Even if the methods of controlling credit expansion had been more efficient than they were, they would have had to be applied at just the right time for best results—and it is surely a matter requiring almost superhuman knowledge to decide just where desirable credit expansion ends and undesirable overexpansion begins. If the time for applying the powers could have been accurately determined, the central banking authorities would still have had to find the courage to apply them at this time without regard for business, political, or international interests which might press for the continuation of an easy money policy. It is sometimes charged that the Federal Reserve Board and reserve banks actually yielded to such influences at critical moments in periods of prosperity and credit expansion.

However ineffective the powers for credit control may have been in periods of prosperity, they were even less effective in stimulating credit expansion and a rebirth of business activity in depressions. When member

banks are doing very little lending, an increase in their reserve accounts at the federal reserve banks by means of open-market operations (in this case, by purchases of government bonds by the federal reserve banks) is not very stimulating, and, when member banks do not need to rediscount, it does not matter if the federal reserve banks lower the rediscount rates. Similarly, when businessmen think they can obtain earnings of zero or minus 10 per cent by borrowing funds at the member banks, they cannot be brought in to borrow by lowering the interest rate charged by member banks from, say, 6 per cent to 4 per cent. However, it is doubtless true that, if expansions of bank credit could be controlled, we should not need to worry very much about periods of contraction.

Changes in Federal Reserve Powers for Credit Control. The inadequacy of the powers of the Federal Reserve System for controlling commercial credit was quite evident by 1933, and many of the provisions of the banking legislation passed by Congress in 1933 and 1935 were directed toward this problem. The Board of Governors now has, for all practical purposes, the actual power to determine the rediscount rates which the federal reserve banks must charge to member banks for advances of funds on the basis of eligible commercial paper. Each federal reserve bank must submit its rediscount rate to the Board of Governors for approval at least once in two weeks, or more often if the Board requires it, and a proposed rediscount rate cannot go into effect unless it is approved. Thus the Board of Governors, by refusing to approve any rate higher than, say, 6 per cent or any rate lower than that figure, can actually determine the rate to be charged. There is little reason to expect that attempts to control credit expansion by changing the rediscount rates alone will be highly effective in any case, but the power is now centralized in one agency and all the federal reserve banks can be made to act together in this matter.

The federal reserve scheme for open-market operations has also been changed. The Federal Open-Market Committee of twelve members formerly acted only in an advisory capacity, but now it has control over the open-market operations of the system. For example, when it decides that government bonds should be sold in the open market by the federal reserve banks, its decision is binding on these banks. The members of the committee formerly represented the twelve federal reserve banks, but now only five of the twelve members represent these banks, the members of the Board of Governors making up the other seven members. This means, of course, that the Board of Governors acting as a unit can control the open-market operations of the Federal Reserve System.

In addition to the strengthening and centralizing of these two old powers for controlling the expansion of bank credit, some important new powers have been given to the Board of Governors. For one thing, the board has the power to change the reserve requirements for the demand

and time deposits of member banks. It may increase these requirements to any amount up to double the basic rates of 13, 10, and 7 per cent for demand deposits in the three classes of member banks, and 3 per cent for time deposits in all classes of member banks. Having once increased the reserve requirements, it may lower them again, but not below the basic reserve requirements just noted. Doubling the reserve requirements behind demand deposits has the rather obvious effect of cutting in half the ability of member banks to make loans resulting in such deposits on the basis of reserves of any given amount held at the federal reserve banks. On several occasions since 1935 the Board of Governors has exercised its power to change the reserve requirements.

Second, the ability of member banks to make loans on the basis of stocks and bonds has been brought under strict control. The Board of Governors has the power to prescribe from time to time for each federal reserve district the percentage of their capital and surplus which member banks may lend on the basis of stocks and bonds. Moreover, under the Securities Exchange Act of 1934, the Board controls margin requirements; that is, it controls the percentage of the market values of securities of various classes which member banks, brokers, and others may advance to borrowers. Third, the Board of Governors may deny member banks the use of the rediscounting and other facilities of the Federal Reserve System as a penalty for making improper use of bank credit for speculative or other purposes which may tend to interfere with the maintenance of sound credit conditions. Formerly, the federal reserve banks had to rediscount eligible commercial paper whenever it was presented by member banks. Closely related is the power to suspend the right of member banks to borrow on their own promissory notes from the federal reserve banks, and to declare outstanding loans due and payable if the member banks persist in undue lending on stocks and bonds. Finally, the Board of Governors may suspend or remove officers and directors of federal reserve banks or member banks for improper activities in the extension of bank credit.

The Adequacy of Federal Reserve Powers for Credit Control. By itself any one of these powers of the Board of Governors might be inadequate to check the expansion of bank credit in a period of prosperity. However, if the Board exercised all or most of its powers at any one time or successively, they should be sufficient to prevent the overexpansion of bank credit under any ordinary conditions. Of course, the problem of applying the powers at the right time still remains, as does that of getting the Board of Governors to use its powers regardless of political or international influences. Up to the present time, the new powers of the Board for controlling the expansion of bank credit have never received a fair trial. The powers had been available only a few years before the United States became involved in World War II, and an emergency situation came

into being in which the federal government was unwilling to rely on these federal reserve powers.

Participation in the war brought about a great increase in the expenditures of the federal government from $12,774,000,000 in the fiscal year 1941 to $78,179,000,000 in 1943, $93,744,000,000 in 1944, and $100,405,000,000 in 1945.[8] This increase in expenditures was financed in part by sharp increases in taxation and by sales of government bonds to individuals and business firms. These methods of government finance are considered noninflationary, since they reduce purchasing power in the hands of individuals and firms while increasing purchasing power in the hands of the government. However, another large part of the increase in governmental expenditures was financed through the sale of government bonds to the banks.

This method of finance is inflationary in character, since individuals and firms do not give up funds as the government receives funds. That is, when banks purchase government bonds, they pay for them by setting up demand deposits for the federal government. When the government spends these deposits created for it by the banks, the funds find their way directly or indirectly into the hands of individuals and firms, and are redeposited in the banks, where they become private, rather than governmental, deposits and can be used by their owners for any productive or consumptive purpose. In such a situation there is little that the Board of Governors of the Federal Reserve System can do to limit the growth of bank deposits. Thus the total deposits of all member banks increased from $61,717,000,000 on December 31, 1941, to $118,378,000,000 on June 30, 1945.[9] Over the same period, the demand deposits of individuals, partnerships, and corporations in the member banks increased from $33,061,000,000 to $57,417,000,000, and the time deposits of individuals, partnerships and corporations in the member banks increased from $11,878,000,000 to $21,254,000,000.[10] For all banks in the United States, demand deposits increased from $44,316,000,000 to $96,730,000,000 and time deposits from $26,476,000,000 to $41,710,000,000,[11] over the same period, while money in circulation in the United States increased from $11,160,000,000 to $26,746,000,000.[12]

Since the method of financing governmental expenditures in wartime by selling bonds to the banks provides the government with funds without taking them away from private individuals and firms, the amount of money income left to individuals and firms in each year, even after paying taxes and buying bonds out of current income, is likely to be far more than enough to buy the commodities and services available for civilian consump-

[8] Press release of the Treasury Department, July 2, 1941, and *Federal Reserve Bulletin*, October, 1945, p. 1047.
[9] *Federal Reserve Bulletin*, October, 1945, p. 1032.
[10] *Ibid.*, p. 1035.
[11] *Ibid.*, p. 1032.
[12] *Ibid.*, p. 1029.

tion at stable prices. If individuals and firms try to spend all their money incomes for the commodities and services left over after the government has satisfied its needs for war in each year, the result is likely to be a serious, and possibly disastrous, inflation of prices. Our government was anxious to avoid such an inflation during World War II, both because rapidly rising prices greatly increase the money cost of a war and because they produce extreme hardships for many classes of citizens. As a result of the desire to restrain inflation, the government ceased depending upon monetary management and banking controls over credit expansion as devices for stabilizing prices, and turned to direct price control. On the whole, the direct control of prices by the government through the Office of Price Administration operated successfully, and it may be some years before this system of price control is entirely abandoned.

SAFETY FOR DEPOSITORS

The Availability of Reserves. Besides providing elasticity of money and commercial credit, and adequate controls over their expansion and contraction, the Federal Reserve System was intended to correct the weakness of the National Banking System with respect to providing safety for depositors. It will be remembered that, under the old system, the reserves held by national banks behind deposits were not available in times of need because of the system of pyramiding or redepositing these reserves in other banks, and that the cash reserves actually held often fell below the supposed minimum amounts because of the system of redepositing reserves and the practice of counting "float" as reserves. Moreover, no system was provided to enable national banks to secure additional reserves in times of need. The practices which caused so much difficulty under the National Banking System were all abolished under the Federal Reserve System. Member banks may keep balances with other member banks, but these balances cannot be counted as reserves. The cash which a member bank keeps on hand is not a part of its legal reserves, nor can it count checks and other items collectible from banks.

The only thing which makes up the legal reserve of a member bank under the Federal Reserve System is its realized balance on account at the federal reserve bank of its district. This balance may be increased by depositing money at the federal reserve bank, by collecting checks through the federal reserve bank and leaving the funds on deposit there, by selling government bonds to the federal reserve bank, by rediscounting or otherwise borrowing at the federal reserve bank, and by other methods. It matters little how the reserve is created. The important fact is that it must be held at the federal reserve bank, where it will always be available when needed. Since all member banks in a district keep their reserves in the federal reserve bank, a large pool of reserves is created in each district.

Obtaining Additional Reserves. The mere availability of fractional reserves behind deposits is not enough to provide safety for depositors. If a bank, in time of need, can have access only to its own reserves, nothing less than a 100 per cent reserve behind deposits can furnish complete safety for depositors. However, fractional reserves behind deposits may be quite adequate if the individual banks can obtain additional reserves when necessary. The principle involved is similar to one of insurance. If a man wants to insure his house against fire, and wishes to function without benefit of insurance company, he needs a reserve equal to the value of the house. However, when he and other homeowners insure their houses, he can be safe by setting aside a small reserve (premium) annually in the hands of an insurance company, confident in the knowledge that if he sustains a loss he can have access not only to the sum which he set aside but also to the funds paid in by homeowners who suffer no losses.

Thus a member bank which holds a reserve of only 13, 10, or 7 per cent behind its deposits is safe if it can obtain additional reserves in times of emergency. Under the Federal Reserve System, if a member bank is hard pressed by depositors who demand their deposits in cash, it can ordinarily obtain additional reserves in cash from its federal reserve bank. To its federal reserve bank it takes eligible commercial paper, received from businessmen in connection with loans which gave rise to deposits, rediscounts this paper, takes the proceeds in federal reserve notes, and goes on to pay off its depositors as long as they ask for their money, which will not be long if they see that the member bank is able to meet their demands. Depositors seldom ask for their entire deposits in cash unless they have some reason to fear that the bank will not be able to meet their requests.

The founders of the Federal Reserve System evidently supposed that commercial banks, as such, would hold as investments only government bonds and other securities readily convertible into cash, and would make almost all of their loans commercial loans of the kind which produce commercial paper eligible for rediscount. If a member commercial bank functioned in this way, there could be no question of the safety of its depositors. In time of need its government bonds could be sold for cash either to its federal reserve bank or on the market, and the commercial paper underlying its loans and discounts could be readily converted into cash by rediscounting, so that depositors could get cash for their deposits on very short notice. Since the federal reserve banks were sure to have ample resources for assisting their member banks, and since funds could even be transferred from one federal reserve bank to another if necessary, it was felt that any lack of safety for depositors of member commercial banks would have to be the fault of the member banks and not of the federal reserve banks or the Federal Reserve System.

Bank Failures. In practice the Federal Reserve System was only moder-

ately successful in providing safety for depositors in member banks during the first twenty years of its operation. Some 995 member banks, with deposits of $492,000,000, failed during the prosperous period from 1921 to 1929, and in the four-year period, 1930–1933, member bank failures and suspensions amounted to 2,110 with deposits of $9,293,000,000.[13] At that, the member banks were much safer than the nonmember banks during these periods, for 4,719 nonmember banks failed in the years from 1921 through 1929, while 6,796 nonmember banks failed or were suspended from 1930 through 1933. It would have been expected that nonmember bank failures would be much more numerous than failures of member banks, since there were many more nonmember than member banks, but, even after due allowance for the difference in numbers is made, the fact remains that depositors in member banks enjoyed much greater safety than those in nonmember banks.

The Causes of Bank Failures. The causes of bank failures in the banking system as a whole were too numerous for us to explore thoroughly. However, in the case of member banks, most failures occurred because the banks were operated in such a way that, under the banking laws, they were not able to claim the assistance of the federal reserve banks. It is sometimes estimated that actual dishonesty played a part in only about 10 per cent of these member bank failures. Of course, when a bank president or cashier "borrowed" a large sum of money from his bank, sometimes without even remembering to leave a promissory note to represent the transaction, and lost the money playing the stock market, there was nothing that the federal reserve bank could do to help the depositors of the bank. Clearly it could not rediscount the promissory note which the bank official forgot to draw up. In other cases, banks, though managed within the law, responded to the twin stimuli of prosperity and intense bank competition by erecting ornate bank buildings which cost more than the entire amount of capital funds furnished by the owners of the banks. When bank buildings were built with funds belonging to depositors, it was difficult to pay off depositors in cash either upon demand or after a few days' notice, and the federal reserve banks could not rediscount bank buildings or parts thereof.

Of the factors which placed many member banks beyond the aid of the federal reserve banks the most important was the change which came over the business of member banks in the 1921–1929 period. Over these years, the total loans and investments of all member banks increased by 48 per cent, their investments alone by 67 per cent, loans on securities by 129 per cent, and loans on real estate by 214 per cent, but their short-term commercial loans which furnish eligible commercial paper did not increase

[13] *Statistical Abstract of the United States*, 1939, p. 263.

at all.[14] In the great depression following 1929, when prices of securities and real estate fell precipitously, many member banks found themselves in a most difficult position. They could not sell their investments for sums large enough to replace the amounts invested. They could not collect their loans based on stocks and bonds, or sell the securities used as collateral for enough to recover the funds which had been turned over to the borrowers. These loans should have been safe if the bankers had confined them to amounts which represented moderate fractions of the current market values of the securities. However, in many cases the banks, acting under the influence of great prosperity and intense competition, made loans on securities which closely approximated in amount the greatly inflated market value of the securities in the boom period, and the great decline which took place in security prices made such loans unsafe.

Real-estate loans could not be collected from the borrowers and, while the banks could become owners of the real estate by foreclosing mortgages, they could not sell the real estate immediately for enough to recover their funds. Finally, since it was not eligible commercial paper, the member banks could not rediscount the commercial paper which resulted from loans on securities or real estate, or use it as security for direct advances from the federal reserve banks. When a dash of hysteria on the part of bank depositors and a desire to withdraw deposits from the banks in cash were added into this mixture, all the ingredients were present for producing bank failures in large numbers even within the Federal Reserve System.

The member banks do not deserve all of the blame for expanding loans based on securities and real estate, instead of short-term, self-liquidating loans to business and industrial firms, during the 1921–1929 period, for they cannot make loans on commercial paper which will later be eligible for rediscount unless such paper is presented to them for discounting. Among the factors which have been used to account for the change which took place in the loans and discounts of member banks are (1) smaller inventory requirements, hand-to-mouth buying, and large profits for business and industrial firms, all of which tended to reduce the needs of these firms for commercial credit; (2) stock market conditions which enabled firms to obtain, by selling stock, funds for working capital which might otherwise have been borrowed from banks; and (3) an "easy money" policy on the part of the Federal Reserve System which placed it squarely up to the member banks to make whatever kinds of loans they could or else hold idle funds. However, the large number of bank failures which took place under the Federal Reserve System indicated clearly that the member banks would have to be compelled to limit themselves to commercial banking or

[14] *Hearings*, Senate Resolution 71, Seventy-first Congress, 3rd Session, p. 138, as reported by Lawrence W. Towle in "Time Deposits and Price Stability," *American Economic Review* December, 1935, pp. 635–666.

else a method would have to be provided to enable them to obtain funds with which to pay off depositors even though they made large numbers of noncommercial loans. Strangely enough, the latter alternative was actually chosen.

Direct Advances to Member Banks. Thus the Banking Act of 1935 provided for a new type of direct loans from federal reserve banks to member banks. Formerly the federal reserve banks could only rediscount eligible commercial paper or make direct advances to member banks for fifteen days or less on the promissory notes of member banks secured by eligible commercial paper or securities of the federal government. Now, when these resources are exhausted, the federal reserve banks are permitted to make direct advances of funds on the promissory notes of member banks, for periods of four months or less, and these promissory notes may be secured *in any way which is satisfactory to the federal reserve banks at the time.* The fact that the interest rate on such advances must be at least one half of one per cent higher than the highest discount rate prevailing at the federal reserve banks at the time is supposed to restrict the use of these loans to emergency periods.

Had this provision existed in earlier years it would probably have saved many member banks from failure, for it enables member banks to convert almost any kind of loan or investment into cash for paying depositors in time of need. Government securities may be sold or borrowed on as usual, eligible commercial paper may be rediscounted, and, under the new provision, any other securities which the member banks may legally own or any commercial paper, ineligible for rediscount because it is based on real estate or stocks and bonds or for any other reason, may be used as security for obtaining these new direct advances, if only the federal reserve banks are willing to accept these assets as security. And, it may be added, if member banks seem doomed to failure unless they receive these new direct advances, the federal reserve banks might be willing to accept almost anything as security for their assistance. Thus if member banks operate within the law and are able to restrain themselves from making large numbers of loans which can never be collected at all, it is difficult to see how member banks can fail.

Other Provisions Affecting Safety for Depositors. The banking laws of 1933 and 1935 contained other provisions which were intended to keep member banks out of trouble and thus indirectly promote the safety of depositors. For example, the Banking Act of 1933 made an effort to separate commercial and investment banking. In the boom years of the 1920's many commercial banks had developed and controlled "investment affiliates" to deal in securities and carry on investment banking activities which they could not carry on legally themselves. The relations of some commercial banks with their investment affiliates left little to be desired, but in

other cases serious abuses crept in. In extreme cases commercial banks would make loans to their affiliates which were exorbitant in amount and based on doubtful security, or they would purchase from their affiliates types of securities which they would not otherwise have considered as good investments. Accordingly the act required all member banks to give up their investment affiliates within a year. Under the same act investment banks are not allowed to hold demand deposits for customers, and no member bank may have as an officer or director any individual who is also an officer or director of an investment banking house. Member banks are not allowed to underwrite security issues, except those of states and municipalities, may not invest more than 10 per cent of their own capital and surplus in the securities of any one company, and may deal in securities only as agents of their customers.

Again, new restrictions have been placed on certain types of loans. Member banks may still make call loans, or day-to-day loans to brokers for the carrying of stocks and bonds, on their own account, but they may not make these loans for corporations, individuals, or foreign interests which have funds to devote temporarily to speculative uses but which may wish to withdraw them on short notice. The Banking Act of 1933 forbade executive officers of member banks to borrow from their own banks and required them to report personal loans received from other banks. The 1935 act, however, permits member banks to make loans to officers in amounts up to $2,500 if the loans are approved in advance by a majority of the directors of the lending banks. Finally, member banks are not allowed to pay any interest on the demand deposits of their customers, and the Board of Governors of the Federal Reserve System has the power to limit the rates of interest which member banks may pay on time or savings deposits.

As a matter of fact, any of the powers recently conferred on the Board of Governors of the Federal Reserve System for controlling the expansion of bank credit might be listed as changes affecting safety, at least indirectly, as well, for surely member banks will operate more safely if they do not overexpand credit. In particular, if loans on stocks and bonds are as dangerous as the writers of recent banking legislation seemed to think, the powers of the Board of Governors to set margin requirements for security loans and to limit the percentage of member banks' capital and surplus which may be tied up in security loans at any one time must be considered as factors affecting the safety of member banks as well as their expansion of credit. Similarly, the power to deny member banks access to the rediscounting and other credit facilities of the Federal Reserve System may be considered to affect their safety as well as their ability to expand credit.

Deposit Insurance. If member banks should persist in failing in spite of everything which has been done to prevent it, we now have a system

of deposit insurance to protect the depositors against loss. The Banking Acts of 1933 and 1935 provided for the establishment of the Federal Deposit Insurance Corporation (which we shall refer to as the FDIC), with its original capital contributed by the federal government and the federal reserve banks. Under this system of deposit insurance, each depositor of an insured bank has his deposits insured up to the amount of $5,000; he may have much more than $5,000 of deposit insurance by maintaining deposits in several insured banks. Although $5,000 may seem a rather low limit on deposit insurance, it was estimated in 1938 that this amount was great enough to provide complete coverage for 98.4 per cent of all depositors' accounts by number, although it covered less than half of all deposits by value.[15]

The insured banks have been paying an annual assessment of one twelfth of 1 per cent on total net deposits for this insurance. When an insured bank fails, the FDIC takes over and either pays off depositors directly or arranges deposits to their credits in other banks. All member banks in the Federal Reserve System must have this deposit insurance, and even nonmember banks are permitted to have it. New national and state member banks must secure deposit insurance before starting out in business, and nonmember banks not already insured must apply to the FDIC and fulfill certain requirements in order to secure deposit insurance. On June 30, 1945, the FDIC was insuring deposits in 13,282 commercial banks, or about 95 per cent of all commercial banks. The demand and time deposits of individuals, partnerships, and corporations in these insured commercial banks amounted to $91,871,000,000 at the time, while their total deposits, including interbank and governmental deposits, amounted to $134,282,000,000.[16]

Now that deposits are insured and member banks can obtain assistance from their federal reserve banks on the basis of almost any type of loan or investment in time of emergency, there can be little question concerning the safety of the depositors of member banks. However, we may note that safety has been secured not so much by improving commercial banking standards, methods, and practices as by making it next to impossible for depositors to lose their money no matter how far the member banks may wander off the reservation of strict commercial banking. Although increased safety for depositors is a welcome development, it would still be possible for one to be dubious about the quality of a commercial banking system which needs devices such as those just described in order to provide safety for depositors.

[15] *Annual Report of the Federal Deposit Insurance Corporation*, 1938, p. 80.

[16] Federal Deposit Insurance Corporation, *Assets and Liabilities of Operating Insured Banks, June 30, 1945*, p. 3.

Commercial Credit. Our discussion of investment and commercial credit has been confined to the economic system of the United States, but we should not suppose that these types of credit are of importance only in capitalistic economic systems. As a matter of fact, the existence of commercial credit is inevitable in any type of economic system which employs modern methods of production. That is, if indirect production is used, it is completely impossible to make all the various expenses which must be met coincide in time with the receipt of payments for the final products. Raw materials and supplies must be purchased in advance of production; expenditures for labor, power, and other things must be made as production proceeds; and some time later payment is received for the finished product. In the meantime, someone must "wait"; that is, someone must extend commercial credit by giving up money or goods at present in return for an equivalent value later. It is no surprise, then, to find a full-fledged system of commercial credit in operation in Soviet Russia, or to learn that such a system was formerly used in the fascist countries. Even under communism, if the use of money could be avoided, it would still be necessary to make efforts and undergo sacrifices long before the resulting satisfactions became available. Thus, in effect, commercial credit would still exist.

Actual methods and practices in connection with short-term credit may differ considerably from one economic system to another. One economy may have a system of commercial banks, while another may not; that is, the government may extend commercial credit directly to productive enterprises, or some productive enterprises may be required to grant such credit to others. The commercial banking system, if there is one, may be privately owned or it may be a governmental monopoly. The persons who run the commercial banks may or may not have control over the total volume of commercial credit issued. Advances of funds may be repayable or nonrepayable. Interest may or may not be charged, and so on, but commercial credit must exist in any modern economic system.

Investment Credit. Differences in methods and practices will also exist in connection with investment credit operations in different types of economic systems. The American corporation may obtain investment credit by selling bonds, which must be paid off with interest, through a private investment banker who is an authority unto himself as to whether the funds should or should not be extended to the corporation. A Russian enterprise may receive a nonrepayable grant of long-term funds, which bear no interest, from a governmental monopoly which extends the funds merely as a part of the economic plan for the whole economy and without any discretion as to whether the funds should be extended or as to the amount which should be given. Underneath these surface differences the simple

fact remains that, in any economic system, there is no way in which factories and machines can be constructed and made to yield all their benefits or products immediately. Someone must advance funds, or economic goods, so that these fixed capital goods may be made, and must wait over a period of years before all the products or satisfactions which the capital goods yield have been obtained. The use of fixed capital necessitates investment credit in any economic system.

QUESTIONS AND PROBLEMS

1. Explain the requirements which a commercial bank must meet if it is to be a member of the Federal Reserve System.
2. Why have many commercial banks remained outside the Federal Reserve System? Explain.
3. Show how the federal reserve banks are owned and controlled, and what functions they perform.
4. Discuss the nature and functions of the Board of Governors of the Federal Reserve System.
5. "Our chief type of money in circulation under the Federal Reserve System is elastic." Do you agree? Explain.
6. "The requirement of eligibility for rediscounted commercial paper is more likely to limit the issue of federal reserve notes than is the requirement of gold (certificate) reserves." Show whether you agree.
7. "Under the Federal Reserve System deposit currency and federal reserve notes depend upon almost exactly the same process for their elasticity." Do you agree? Explain.
8. "On the basis of reserves obtained by rediscounting, a single member bank may not be able to expand demand deposits to anything like the legal limit." Show whether you agree.
9. "Although both are elastic, demand deposits in member banks are much more expansible than federal reserve notes." Explain.
10. "In the period from 1913 to 1933, federal reserve notes and deposit currency fulfilled completely the requirements of elasticity." Do you agree? Explain.
11. How were changes in rediscount rates supposed to control the expansion and contraction of federal reserve notes and deposit currency? Show why this method of control was ineffective.
12. "Open-market operations were likely to be no more and no less effective than rediscount rate changes in controlling the expansion of deposit currency." Show whether you agree.
13. "Credit control powers were even less effective in depression than in prosperity." Explain.
14. How were the powers for credit control involving rediscount rate changes and open-market operations strengthened under recent banking legislation? Explain.
15. "Many important new powers for controlling the expansion of deposit currency were provided by the Banking Acts of 1933 and 1935." Explain.

16. "In the emergency created by World War II, the government resorted to direct price control in order to prevent inflation instead of relying on the powers of the Federal Reserve System operating through credit control." Explain.

17. "Although reserves against deposits were smaller under the Federal Reserve System than under the National Banking System, deposits were much safer under the former system even before 1933 than they were under the latter." Show whether you agree.

18. "Reserves behind deposits were centralized under the Federal Reserve System and a method was provided which enabled member banks to obtain additional reserves in time of need." Explain.

19. "Even before the banking legislation of 1933 and 1935, the depositors of any soundly operated member bank were safe at all times under the Federal Reserve System." Discuss.

20. Compare the record of member and nonmember banks with respect to safety in the years from 1921 to 1933.

21. "Most member bank failures prior to 1933 resulted from imprudent rather than dishonest management." Explain.

22. "In the years from 1921 to 1929, many member banks placed themselves in such a position that the federal reserve banks could not give them adequate assistance in an emergency." Explain.

23. "Although the rediscounting process remained unchanged, the Banking Act of 1935 greatly increased the ability of member banks to obtain assistance at the federal reserve banks in times of need." Do you agree? Explain.

24. Discuss the system of deposit insurance which is in effect at the present time.

25. "Deposit insurance and the system of direct advances from federal reserve banks to member banks may provide increased safety for depositors, but these devices would never be necessary in a soundly operated commercial banking system." Show whether you agree.

26. "Most of the powers of the Federal Reserve System for controlling the expansion of bank credit may also be listed as factors affecting safety for depositors." Explain.

27. "The use of commercial and investment credit cannot be avoided in any type of economic system, however much methods and practices in connection with the use of these types of credit may differ from one system to another." Show whether you agree.

REFERENCES FOR FURTHER READING

Bradford, F. A., *Money and Banking.* New York: Longmans, Green and Company, rev. ed., 1941, Chaps. 5–10, 12–14, 17–24.

Burgess, W. R., *The Reserve Banks and the Money Market.* New York: Harper & Brothers, 1936, Chaps. 1–8 and 15–18.

Clark, L. E., *Central Banking under the Federal Reserve System.* New York: The Macmillan Company, 1935, Chaps. 1, 2, 7–13, 19.

Dewing, A. S., *The Financial Policy of Corporations.* New York: The Ronald Press Company, 1934, Book III, Chap. 2, and Book V, Chaps. 8–12.

Kemmerer, E. W., *The A B C of the Federal Reserve System*. Princeton, N. J.: Princeton University Press, 10th ed., 1936, Chaps. 2, 3, 6, 7, and 10–14.

Peterson, J. M., Cawthorne, D. R., and Lohman, P. H., *Money and Banking*. New York: The Macmillan Company, 1941, Chaps. 5–14, 25–28.

Thomas, R. G., *Our Modern Banking and Monetary System*. New York: Prentice-Hall, Inc., 1942, Chaps. 4–22, 27.

Westerfield, R. B., *Money, Credit, and Banking*. New York: The Ronald Press Company, 1938, Chaps. 9, 11–21, 30–37, 39, 41, and 46–48.

XXVI

Foreign Exchange

We have examined at some length the means which are employed in carrying on domestic trade and in settling domestic obligations, and it is only a short step to the study of the ways in which obligations are settled between different countries. International trade is similar to domestic trade in many respects, but methods must be found of getting around certain difficulties in international trade which do not affect domestic trade. One of these difficulties lies in the fact that the various trading countries have different monetary units and systems, so that the ordinary money of one country is not generally acceptable in other countries.

Of course, the gold money which a country uses, if any, would be accepted in settlement of obligations in other countries, but only as so much gold by weight and fineness and not as so many units of money. However, there are several factors which make gold rather unsatisfactory as a medium for settling international obligations. Much of the world's gold supply is ordinarily tied up in individual countries as reserves behind paper money or bank deposits and hence is not free to be used for making international payments. If all international obligations had to be settled by means of gold shipments, the gold readily available for the purpose would have to move about quite rapidly in years in which the total of international trade reached 60 or 70 billions of dollars by value. Moreover, the distribution of the total stock of gold among countries may not be well adapted to their respective volumes of international trade. Finally, gold is rather expensive to ship, there is a certain loss from abrasion when gold is repeatedly handled, and considerable amounts are sometimes lost in transit, as the quantities of gold which decorate the ocean floors testify.

THE USES OF FOREIGN EXCHANGE

Under these conditions it is not surprising that a method has been developed which makes it possible to effect international payments without the use of very much gold. Most international obligations are settled through foreign exchange, or the balances which bankers in individual

countries carry with banks in other countries. On the basis of these balances, credit instruments called bills of exchange are used in making most payments between countries. The bill of exchange is not really a new type of credit instrument at all, but is merely our old friend the draft, which becomes dignified by a more pretentious title when it is used in international transactions.

Financing an Export Transaction. In illustrating the use of bills of exchange, let us suppose that a certain exporter in New York has sold a quantity of merchandise to an importer in England and has quoted his customer a price of £1,000 for the goods. When the exporter ships the goods, he draws a bill of exchange on the importer, ordering the importer to pay £1,000 to the exporter either at sight (when presented) or after a certain number of days, according to the terms of the transaction. The exporter takes this bill of exchange to a bank in his city, and the bank buys it from him at whatever the going price is on that day for bills of exchange of this type. The American bank sends the bill of exchange over to a correspondent bank in England for collection. The correspondent bank may be merely a foreign branch of the American bank or it may be an independent bank which has entered into an agreement with the American bank under which the two banks agree to pay and collect bills of exchange and perform certain other functions for each other. The correspondent bank, upon receiving the bill of exchange, presents it to the English importer, who pays the sum called for at once if the bill is payable at sight, or accepts the bill for payment after the designated number of days if it is a time bill. In this way the American bank's supply of foreign exchange (or its balance with the English correspondent bank) is increased.

Conveniences of Using Foreign Exchange. The convenience of using bills of exchange to finance such export transactions is worthy of brief comment. If we asked the exporter how he would like to be paid for his exported merchandise, he would probably express a preference for being paid at once, at home in his own country and city, and in dollars; and these are the results achieved by using the bill of exchange. The exporter receives his payment as soon as the merchandise is gone, he is paid by a bank in his home city, and he receives dollars rather than pounds sterling. Moreover, this convenience for the exporter is not obtained at the expense of the importer, for the importer makes payment for the goods in his home town, in the money of his country (pounds sterling), and only when the merchandise has arrived and is available for his use.

Safety for the Bank. While the conveniences of using bills of exchange are easily seen, students often wonder how a bank dares to buy a bill of exchange from an exporter for cash. How does the bank know that the goods have been shipped, that there is such an English firm as the one named in the bill, or that this firm will make payment when the bill reaches

England? Ordinarily the bank's position in buying such a bill of exchange is not very precarious. The exporter, upon turning the goods over to a steamship line for transportation to England, receives what is called an ocean bill of lading. This document acknowledges receipt of the goods and also carries title to them. Although it is made out to the exporter, it will be endorsed by him in blank and turned over to the bank, so that the bank has title to the goods while the bill of exchange is in the process of collection. The exporter also gives the bank a certificate of marine insurance covering the goods while they are in transit and, among other documents, an hypothecation paper which gives the bank the right to sell the exported goods as best it can if payment of the bill of exchange is refused. These various documents are sent along with the bill of exchange to the correspondent bank in England. The importer must pay the bill of exchange, or agree to pay it when due, in order to get the bill of lading and obtain title to the merchandise. Thus we see that the bank is rather well protected in its purchase of the bill of exchange from the exporter.

Financing an Import Transaction. In the transaction just examined, the importer made payment and the exporter received payment, but the American bank gave up money or granted a deposit to the exporter in the United States and received in return only a credit to its account at the correspondent bank in England. The American bank gets its money back by selling other bills of exchange, drawn against its account with the correspondent bank in England, to American importers or other persons who wish to make payments in England. By selling these bills of exchange to importers or others at a rate of exchange or price which is somewhat higher than that which the bank is currently paying to exporters for other bills, the bank is able to derive a certain margin of income from its transactions in foreign exchange.

When an American importer has purchased a quantity of merchandise from an English exporter and has agreed to pay him £500 for the goods, the importer will come to the bank and purchase a bill of exchange for this amount payable in England, giving the bank either money or a check against a demand deposit owned by the importer. This bill of exchange is a bank bill, because it is drawn by the American bank against its correspondent bank in England. That sold by the exporter was a trade bill since it was drawn on an English business firm. The importer sends the bill of exchange to the exporter in England, who presents it at the correspondent bank and receives payment in pounds sterling. The result is a reduction in the American bank's supply of foreign exchange, or its balance with the English correspondent bank.

It may be noted that either of these transactions which we have described could have been financed from the opposite direction. In the case of the export transaction, the English importer could have purchased from

his bank a bill of exchange drawn in dollars and payable in the United States. This bill would then have been sent to the American exporter, who would convert it into cash by presenting it at the English bank's correspondent bank in this country. Again, in the case of the import transaction, the English exporter could have drawn a bill of exchange in terms of dollars on his customer in this country, and obtained cash by selling this bill to his bank, which would have collected it from the American importer through its correspondent bank in this country. Whether the importer or the exporter is to take the initiative in the settling of an international obligation is a matter for the parties to the transaction to decide.

Triangular Exchange. In our two transactions the bill of exchange purchased by the importer only half offset the bill of exchange which the exporter sold to the bank. This should not be a matter of any great concern for, if the bank in question buys and sells many bills of exchange daily or weekly, it should be able to make its purchases and sales balance quite closely. However, if our trade with England is temporarily unbalanced and we are exporting many more goods to England than we are importing from that country, our bank, along with other banks, will find that it is purchasing many more bills from exporters than it is selling to importers and others. In this case the banks' balances with their correspondent banks in England will become unduly large, and the exchange rates on England will go to a discount in this country to indicate the unwillingness of the banks to buy more bills of exchange payable in England from exporters and their strong desire to sell larger numbers of bills to persons and firms who desire to make payments in England. If the banks' balances with their correspondent banks in England remain too large in spite of the downward movement of the exchange rate in this country and if the two countries are on the gold standard, the American banks may require their correspondents to ship a quantity of gold in order to reduce their balances in England.

However, even in this situation little or no gold may actually move from England to the United States if our banks can sell bills of exchange drawn on their balances in England to American importers who have to make payments in some third country, such as, say, Brazil. Suppose, for example, that our exports to England in a certain period amount to $400,000,000, while our imports from that country are only $300,000,000. Since $100,000,000 of these exports are not canceled by imports from the same country, our banks find that their balances with correspondent banks in England are increased by that amount after the bills of exchange necessary to finance this trade have been bought and sold. Suppose next that American exports to Brazil in the same period were valued at $100,000,000, while our imports from Brazil amounted to $200,000,000. The exports to

Brazil set up the necessary balances for paying for half of our imports from that country, but how will our importers pay for the remainder of the imports? The answer is that banks in the United States will sell these importers bills of exchange, drawn against their balances with correspondent banks in England, to the extent of $100,000,000. These bills of exchange will be sent to the Brazilian exporters who will sell them to the Brazilian banks, which in turn will send them to correspondent banks in England for collection from the correspondent banks of the American banks. Thus the excessive balances of our banks with correspondent banks in England are reduced.

But does not all this merely mean that gold will have to flow from England to Brazil instead of from England to the United States? Not necessarily. Brazil usually imports more from England than she exports to England, and we may assume that, in the period in question, Brazilian imports from England amounted to $300,000,000 while her exports to England were valued at only $200,000,000. In this case the exports furnish the direct means of paying for two thirds of the imports. To the importers whose imports were not covered by exports directly to England the Brazilian banks will sell bills of exchange drawn against the balances in correspondent banks in England which were set up as the result of Brazilian trade transactions with the United States. These balances just suffice to pay for the Brazilian excess of imports from England, and all three countries are in the clear. All importers have made payment, all exporters have received payment, and no banks have excessive balances remaining with their correspondent banks in any of the three countries.

Our conclusion is that net balances between two countries as a result of international trade, even though they are large, do not necessarily give rise to gold movements of equal magnitude between the countries. Part, if not all, of such net balances will ordinarily be used up in settling balances between the countries under discussion and the various other countries with which they trade. Of course, it would be little short of miraculous if precisely the same net balance were owed, as a result of international trade, between England and the United States, the United States and Brazil, and Brazil and England, in any given period. However, since every export from one country creates an import for some other country, it follows that total exports must equal total imports when all the trading countries are considered, and the chance that any given country will need to make or receive only relatively small gold payments, in carrying on a given volume of international trade, improves as the number of trading countries increases. Just as the sums which individual banks have to pay to or receive from a local clearinghouse are small relative to the total volume of checks cleared, so the net balances which have to be settled in gold between many coun-

tries operating on the gold standard are small in relation to the total volume of international trade.[1]

RATES OF EXCHANGE BETWEEN COUNTRIES ON THE GOLD STANDARD

We have seen that bills of exchange are bought and sold by the banks at certain rates of exchange, but nothing of consequence has yet been said about these rates of exchange and the forces which determine them. Actually several rates of exchange ordinarily prevail at any given time between two countries such as the United States and England. The highest rate of exchange is that for cables or telegraphic transfers, which are orders by cable to pay money abroad. These "cables" are not really bills of exchange, but they have a similar effect and are drawn almost entirely by banks against their foreign branches and correspondents. The high rate of exchange for cables is a result of the speed with which payments may be made by this method. Somewhat lower at any given time is the rate of exchange for sight or demand bills of exchange. These bills are payable when presented in the country on which they are drawn, which means in effect that they are payable after the elapse of the several days necessary to convey the bills to the other country. The rates of exchange for time bills of exchange, or bills which are payable some 30, 60, 90, or 120 days after they are drawn or after they are presented in the other country, are still lower. For our purpose the rate of exchange for sight bills will be considered as basic and the term "rate of exchange," as used hereafter, will be understood to mean this particular rate.

A rate of exchange, from our point of view, is the price in this country of one unit of some other country's money payable in the other country. For example, the rate of exchange between the United States and England at any given time is the price in dollars of a bill of exchange calling for the payment of one pound sterling in England. Since such a rate of exchange is a price, it is determined in much the same fashion as other prices —that is, by conditions of demand and supply. However, before examining the effect of demand and supply conditions on rates of exchange, we must pay attention to some special conditions which prevail in the determination of this particular kind of price. The description of the determination of exchange rates differs somewhat according to whether the countries in question are on or off the gold standard, and we start with the assumption that they are on the gold standard.

The Gold Par of Exchange. Between such countries the rate of exchange varies around a certain norm or basic level known as the gold par of exchange. This par is found by dividing the pure gold content of the

[1] Not all gold which actually passes between countries is being used to settle net balances of trade. Gold also moves from one country to another for safekeeping and is bought and sold as a commodity between countries.

one country's monetary unit by the pure gold content of the other country's monetary unit. Thus if the United States and England were on the gold standard and the pound sterling contained 113.0016 grains of pure gold as against 13.7137 for the dollar, the gold par of exchange would be about $8.24 per pound sterling.[2] Actual rates of exchange do not often equal the gold par of exchange. Par means here, as in golf, a kind of ideal figure which the actual results obtained will only approximate.

The Gold Shipping Points. Although the rate of exchange varies around the gold par of exchange between gold standard countries, the variations are confined within certain limits known as the gold shipping points. Bills of exchange, rather than gold, are used in making international payments because it is cheaper to use them than to use gold. If the cost of using bills of exchange became so great that it was cheaper to make payments with gold than with bills, gold would move between two gold standard countries. Thus the rate of exchange between such countries will never be higher or lower than the gold par of exchange by more than the cost of shipping gold.

Just how great the cost of shipping gold is between, say, the United States and England cannot be stated once and for all. The cost is composed of such items as freight, insurance, loss of interest while the gold is traveling between countries, packing, handling, brokerage fees, and one thing and another. All or most of these costs are able to vary from one time to another and, as they change, the gold shipping points move closer to or farther away from the gold par of exchange. However, if the gold par of exchange between England and the United States is $8.24 to the pound sterling and the cost of shipping one pound sterling in gold in either direction is assumed arbitrarily to be five cents, then the actual rate of exchange can vary only within the limits of $8.29 and $8.19 per pound sterling.

That is, under these conditions, if an exporter in the United States came into a bank to sell a bill of exchange for £1,000 drawn on his customer in England and the bank offered him $8.10 per pound sterling for the bill, or $8,100 altogether, the exporter would refuse to sell. He could direct the customer to have £1,000 in gold sent over to this country and would realize $8,240, less $50 for the cost of shipping, or $8,190 altogether. Thus the rate of $8.19 per pound sterling would be the gold importing point, or the point where it became cheaper to collect payments from England in gold than to use bills of exchange. In the opposite case, an American importer who had to make a payment of £1,000 in England would refuse to buy a bill of exchange from a bank at $8.35 per pound sterling, or $8,350 altogether, because he could buy the necessary amount of gold for $8,240, and, after paying the cost of shipping the gold to England, would be able to make his payment for a total cost of $8,290. The rate of $8.29 per pound

[2] Actually $8.2397.

sterling is therefore the gold exporting point, or the point where it becomes cheaper to make payments in England by sending gold than by using bills of exchange.

In this illustration we have been making the contrary-to-fact assumption that an individual importer or exporter can ship gold in small quantities as cheaply as the banks can ship it in large quantities. But this need not concern us for, in actual practice, importers and exporters are not called upon to ship gold, nor will they encounter a rate of exchange at the banks which is higher than the gold exporting point or lower than the gold importing point. If the rate of exchange falls to the gold importing point, banks will continue to buy (under the assumed cost of shipping gold) at $8.19 to the pound sterling all the bills of exchange which exporters have to sell, and will reduce their excessive balances in England by having their correspondent banks send gold to this country. Conversely, if the rate of exchange rises to the gold exporting point, the banks will continue to sell at $8.29 to the pound sterling all bills of exchange which importers and others want to buy, and will build up their greatly reduced balances with correspondent banks in England by shipping necessary quantities of gold to that country. The banks, in other words, are specialists in the matter of shipping gold. They have the knowledge and facilities which enable them to ship gold more cheaply and efficiently than individual importers and exporters. Thus the banks take care of actual gold shipments, and the rate of exchange between two gold standard countries will vary only within the limits set by the gold shipping points above and below the gold par of exchange.

The Determination of Exchange Rates. We must now consider the demand and supply forces which cause the rate of exchange to fluctuate around the gold par of exchange and within the limits set by the gold shipping points. If our imports and exports with respect to countries other than England are in balance, and in a certain period of time we have been exporting to England more heavily than we have been importing from her, many exporters have been coming to American banks to sell bills of exchange drawn on English importers, while comparatively few American importers have been buying bills of exchange from our banks in order to make payments in England. As a result our banks have excessively large balances with their correspondent banks in England, will be reluctant to increase these balances further, and will be anxious to reduce them. The rate of exchange on England will be at a discount or below par in this situation.

That is, if still more exporters come in to American banks to sell bills of exchange which, when collected, will have the effect of building up still further the banks' already excessive balances with correspondent banks in England, they will be offered a low rate of exchange for these bills. On the

other hand, any importers who come in to buy bills of exchange which will help the banks to reduce their large balances with correspondent banks in England will be met with open arms and will be offered a low rate of exchange. This means that both the banks' buying rate and their selling rate will be at a discount at the same time. If the exchange rate at which the banks are buying bills of exchange from exporters is down around the gold importing point, the rate at which they sell bills to importers will also be at a considerable discount, though somewhat higher than the buying rate. The banks do not make their income from dealing in foreign exchange by buying bills at a discount and selling them at a premium at the same time, although that is what students sometimes think.

In the converse case, with the rest of our trade in balance, our imports from England have been heavy, our exports to England have been light, many importers have been coming into American banks to purchase bills of exchange payable in England, and comparatively few exporters have been approaching the banks to sell bills of exchange payable in England. The banks here find that their balances with correspondent banks in England have been well-nigh exhausted, they are reluctant to sell bills of exchange which will further deplete these balances, and they are anxious to engage in transactions which will build up these balances once more. In this situation the rate of exchange on England will tend to be at a premium or above par in the United States. Any importers who now attempt to buy bills of exchange, which must be drawn against the American banks' already depleted balances with correspondent banks in England, will be asked to pay a high rate of exchange. On the other hand, any exporters who wish to sell bills of exchange payable in England, which will help our banks to build up their balances in that country, will be warmly greeted and will be paid a high rate of exchange, though it will not be quite as high as that which the banks are charging to importers. Thus, in this situation, both the banks' buying rate and their selling rate will be at a premium, with the latter rate somewhat higher than the former.

If pounds sterling are at a discount in the United States, dollars will be at a premium in England, and dollars will be at a discount in England when pounds sterling are at a premium in the United States. This does not mean that, if the rate of exchange is $8.22 to the pound sterling in the United States, it will be $8.26 to the pound sterling in England, or vice versa. The rate of exchange between pounds and dollars will be the same in both countries. Exactly the same rate of exchange means that pounds sterling are at a discount in the United States and dollars are at a premium in England, or vice versa. That is, if the rate of exchange is $8.22 to the pound sterling, or below par from our point of view, people in this country have to pay fewer dollars than usual to get a pound sterling, and pounds sterling are at a discount here. Under the same exchange rate the people

in England get fewer dollars than usual for each pound sterling, and dollars are at a premium there.

Corrective Influences. If the rates of exchange on other countries are at a premium or at a discount in the United States, what will cause these rates to move back toward par? Since heavy exports and light imports on the part of the United States bring rates of exchange on other countries to a discount here, our imports must increase and our exports decrease if the exchange rates are to move upward toward par. In the opposite case, when our heavy imports and light exports bring exchange rates on other countries to a premium in this country, our imports must decrease and our exports increase if the exchange rates are to move downward toward par.

When the trading countries are on the gold standard, there are two factors which help to bring exchange rates back toward par when they have fallen to a discount or have risen to a premium. One of these factors is the influence of the high or low exchange rates themselves. If rates of exchange on other countries are at a discount in the United States because of our previous heavy exports and light imports, these low rates of exchange will tend to discourage exports and stimulate imports. Exporters can charge no higher prices than before to their customers in other countries and, when they try to convert the foreign currencies into dollars, they find that they get fewer dollars than usual for the monetary units of the other countries. Importers have to pay no higher prices than before for foreign merchandise and for fewer dollars than usual they can obtain bills of exchange payable in the currencies of other countries. Thus our formerly large exports decrease, our formerly small imports increase, and exchange rates on other countries tend to rise toward par in this country.

Conversely, if rates of exchange on other countries are at a premium in the United States because of our previous large imports and small exports, these high rates of exchange will discourage imports and encourage exports. Exporters receive more dollars than usual when they convert the prices, charged to customers in terms of foreign currencies, back into dollars, and importers must pay more dollars than usual to get bills of exchange payable in terms of foreign currencies. Thus our formerly small exports increase, our formerly large imports decrease, and exchange rates on other countries fall back toward par in this country. However, since premiums and discounts above and below par cannot be very large between countries on the gold standard, the effects of the high and low exchange rates themselves on imports and exports cannot be very severe.

The second corrective factor is found in gold movements and price level changes amoung countries on the gold standard. If rates of exchange on other countries are falling in the United States because of our large exports and small imports, the movement may continue until the rates of exchange reach the gold importing points. Then gold moves from other

countries to the United States. The total quantity of money in circulation tends to increase in the United States and decrease in the countries which send us gold. The general price level rises in the United States and falls in the other countries. The United States with its high prices becomes a favorable market for the goods of other countries, but a poor place in which to buy goods. The other countries, with their lowered prices, are favorable sources of imports for the United States but poor places in which to sell our exports. Thus our formerly large exports decline, our formerly small imports increase, and rates of exchange on other countries move upward toward par in the United States.

This whole process is reversed under the opposite circumstances. If the United States has large imports and small exports for a time, rates of exchange on other countries will rise here and may reach the gold exporting points. Gold moves from this country to other countries. The total quantity of money and credit in circulation decreases in the United States and increases in other countries. Our price level falls and those of other countries rise. The United States with its lowered prices is able to sell many goods to other countries whose price levels have risen, but it does not buy as many goods as formerly from these countries. Thus our formerly large imports decrease, our formerly small exports increase, and the rates of exchange on other countries move downward toward par in the United States.

Qualifications. The above process has been described without qualification in order that the forces at work might be seen clearly, but the student should be cautioned against believing that it always works automatically and precisely in practice. In examining some of the factors which may interfere with the operation of this gold flow-price level-trade mechanism, let us return to the situation in which large exports and small imports on the part of the United States have brought exchange rates on other countries to a considerable discount in this country so that an inflow of gold would be expected. Such a flow of gold may be forestalled for a time if banks in this country grant credits to banks in other countries and thus build up the depleted balances of foreign banks to some extent. A delay may also be accomplished if gold in other countries is merely "earmarked" for the United States instead of actually being sent here. If gold finally comes to the United States, the effect even under the gold standard is merely to increase the total quantity of money and credit which *can be* in circulation, and not necessarily the total amount of money and credit which really *is* in circulation. Our price level will rise only if increased money and credit really get into circulation, and perhaps not then, for an increased quantity of money and credit in use may be offset by an increased volume of trade within the country or by a decreased velocity of circulation or rate of turnover of money and credit.

The other countries which ship gold to the United States may or may not experience declining price levels. If they had excess gold reserves in the first place, they may have enough gold left after sending us some so that the quantity of money and credit previously in circulation in these countries can be maintained. Finally, even if the price level rises in the United States and falls in other countries, there is some question as to the extent to which our imports will increase and our exports will decrease. Goods are far from being completely free to move between countries, and tariffs, quotas, and other interferences with international trade may hamper the changes in imports and exports which would be expected in the absence of such restrictions. Moreover, the changes in imports and exports depend to some extent on how badly other countries need the products which they obtain from us as compared with the strength or degree of elasticity of our demand for their products. This may not be an exhaustive list of the factors which may condition or interfere with the process in question, but it gives an idea of the difficulties which may be involved in its operation.

RATES OF EXCHANGE BETWEEN COUNTRIES OFF THE GOLD STANDARD

The Purchasing-Power Par. The analysis of rates of exchange is altered considerably when the various trading countries are off the gold standard. In the first place, it is clear that gold pars of exchange are no longer of significance. For example, it is still possible to divide the 113.0016 grains of pure gold which are nominally in the pound sterling by the 13.7137 grains of pure gold which are nominally in the dollar, and get the answer as before, or about $8.24 to the pound sterling, but, if gold cannot be exported freely from England and the United States, this gold par will not be effective in influencing actual rates of exchange between the two countries.

When two trading countries are off the gold standard, the gold par of exchange between them is replaced by another which is called the "purchasing-power par." This new par, as the name implies, depends on the respective domestic purchasing powers of the countries' monetary units in terms of goods capable of moving in international trade. For example, if the pound sterling will buy in England on the average the same quantity of such goods as five dollars will buy in the United States, the purchasing-power par is five dollars to the pound sterling. Since they can no longer require English banks to send gold to this country, our banks are willing to buy bills of exchange payable in England only in order to become able to sell other bills of exchange payable in England. The price which banks in the United States are willing to pay to exporters for bills of exchange drawn on England depends therefore on the price which the banks can charge and which importers are willing to pay for bills of exchange payable in England. Since importers wish to buy such bills only for the purpose of making purchases in England, the rate of exchange which they

are willing to pay depends upon what the pound sterling will buy in England in terms of exportable goods as compared with what the dollar will buy in the United States. Hence the par of exchange depends upon the respective domestic purchasing powers of the pound sterling and the dollar in terms of goods which move in international trade.

The purchasing-power par differs from the gold par in some ways. For one thing, it is a kind of natural par and not one which is officially established and recognized by the governments and banks of the two countries. Second, unless changes in prices in one country are offset by similar changes in the other country, the purchasing-power par is a fluctuating par and is affected by changes in the price levels and purchasing power of money in the two countries in terms of goods which move in international trade. The gold par of exchange between two countries, on the other hand, remains fixed as long as these countries do not change the pure gold content of their monetary units. However, the purchasing-power par is like the gold par in that it does not actually determine the rate of exchange between the two countries but constitutes only a normal level around which the rate of exchange will fluctuate under the influence of demand and supply conditions.

The Elimination of Gold Shipping Points. The variations of the rate of exchange around the purchasing-power par are not confined within narrow limits by a pair of gold shipping points because, if the countries which are off the gold standard do not permit the exportation of gold, there are no points at which it becomes cheaper to send gold between the countries than to use bills of exchange. If the purchasing-power par is $5.00 to the pound sterling, an exporter may go to his bank in the United States to sell a bill of exchange drawn on his customer in England and find that the rate of exchange is actually $4.50 or $4.00 to the pound sterling on that day. Whatever the rate is, he must take it or leave it. He cannot swell with indignation while asserting that he will have his English customer send him gold rather than accept such a ridiculously low rate of exchange. Similarly, an importer who wishes to make a payment in England must pay the going rate of exchange since (by assumption) he does not have the alternative of sending gold to the English exporter.

The Determination of Exchange Rates. While the gold par of exchange and the gold shipping points disappear when two trading countries are off the gold standard, the fluctuations of the rate of exchange around the purchasing-power par are subject to the usual influences of demand and supply conditions. If our trade with other countries is in balance, but the United States has a heavy surplus of exports in trading with England, many exporters come to banks in the United States to sell bills of exchange drawn on England, but comparatively few importers come to the banks to buy bills of exchange payable in England. Our banks' balances with

correspondent banks in England grow rapidly, and, since our banks cannot recover these balances by having gold shipped to this country, they become reluctant to increase these balances further. Thus when still more exporters attempt to sell our banks bills of exchange drawn on England, they are offered a rate of exchange which is very low and well below the purchasing-power par. At the same time importers who wish to buy bills of exchange payable in England, which will help our banks to reduce their swollen balances in that country, are also granted a low rate of exchange but one which is slightly higher than that which is currently paid to exporters. In this situation, then, both the buying rate and the selling rate are at a considerable discount. It will not be necessary to go into the opposite situation in detail, but the rate of exchange on England will tend to be at a considerable premium above the purchasing-power par in the United States, other things being equal, when our imports from England have been large and our exports to that country have been small.

Corrective Influences. When trading countries are off the gold standard, it is clear that we cannot depend upon gold movements between the countries, changes in the quantities of money and credit in circulation in the countries, changes in the price levels of the countries, and, finally, changes in the imports and exports of the countries, to bring the rate of exchange back toward par after it has gone to a premium or a discount. With this process eliminated, the corrective influence, which is of minor importance under gold standard conditions, assumes major proportions when the countries are off the gold standard. This influence is found in the high or low rates of exchange themselves.

Suppose, for example, that the rate of exchange between the United States and England was five dollars to the pound sterling not so long ago, but now, because of large exports and small imports on the part of this country, is four dollars to the pound sterling. Prices in the countries have not changed, the purchasing-power par is the same as before, and our exporters who sell goods in England cannot charge their customers higher prices than before, but when they try to convert the prices charged in pounds sterling back into dollars they receive only four fifths as many dollars as formerly for each pound sterling. The effect on exporters is the same as if the prices obtainable for their goods had declined by 20 per cent, and exports to England are discouraged. On the other hand, imports to this country from England are stimulated. The prices which our importers must pay for goods in England are the same as before, but for 80 per cent as many dollars as formerly they can obtain the bills of exchange drawn in pounds sterling which are necessary to pay these prices. The effect on importers is the same as if the prices of English goods had declined by 20 per cent. Thus our formerly large exports to England decline, our formerly

small imports increase, and the rate of exchange on England rises in this country.

In the opposite situation, when our imports from England have been large and our exports to England small, with the rest of our trade in balance, the rate of exchange on England will be at a considerable premium above the purchasing power par in this country. The high rate of exchange will stimulate our exports to England, since the prices charged by our exporters to importers in England in terms of pounds sterling can be converted into an unusually large number of dollars by selling bills of exchange in this country. Imports from England will be discouraged, since our importers must pay an unusually large number of dollars to obtain the pounds sterling bills necessary to pay the prices charged by English exporters for goods. Thus our formerly large imports from England decrease, our formerly small exports to England increase, and the rate of exchange falls back toward the level of the purchasing-power par. Under ordinary conditions, changes in the exchange rate itself are adequate, through their effect on imports and exports, to prevent a premium or a discount from existing continuously even when countries are off the gold standard.

Governmental Control over Exchange Rates. This discussion of the determination of the rate of exchange between two countries which are off the gold standard has tacitly assumed that, under these conditions as under gold standard conditions, the rate of exchange is allowed to find its own level without governmental interference. This is quite likely not to be true in a practical situation. For example, suppose that England is at war and needs her imports from the United States and other countries very badly and in large quantities, though she is unable, because of wartime conditions, to supply the United States and other countries with the usual volume of exports. If England imports heavily from the United States and exports little to us, dollar exchange will be at a premium in England and pounds sterling exchange will be at a discount in the United States.

Under ordinary conditions this situation with regard to exchange rates would cause English exports to the United States to increase and English imports from the United States to decrease, but we are not dealing here with ordinary conditions. England cannot afford to reduce her imports from the United States and is unable to increase her exports to this country. If the trade goes on in an unbalanced condition, dollar exchange in England will be subject to an ever-increasing premium, and pounds sterling exchange in the United States will sell at an ever-increasing discount. The effect will be that England can obtain her vital imports from the United States only at a continually increasing cost in terms of pounds sterling. Under such conditions England is likely to take drastic steps to control the exchange rate between pounds sterling and dollars.

Such control might be achieved in part by the setting up of a stabiliza-

tion fund in the hands of the English government which could be used to buy and sell foreign exchange in such a way as to keep exchange rates within reasonable bounds. That is, if the English government, through financial agents in this country, offered to buy bills of exchange drawn in pounds sterling from our exporters at a price of four dollars to the pound sterling, the rate of exchange which our bankers would offer to exporters for these bills of exchange would not be expected to fall below that figure. On the other hand, the English government's agents would be instructed to sell part of their holdings of such exchange if the market rate went much above four dollars to the pound sterling. England might take quite a loss in this effort to stabilize the exchange rate, but this loss might well be less than that which would be incurred if the rate of exchange were allowed to fall indefinitely. Besides, by controlling the exchange rate, England would be more certain of her ability to continue to receive the necessary imports from the United States.

England attempted direct control over the exchange rate between pounds sterling and dollars in both World Wars I and II. In addition, other steps were taken to relieve the foreign exchange situation. England restricted the use of her limited supplies of foreign exchange for importing various goods which were deemed nonessential, in order to preserve foreign exchange for the importation of important war materials. Before the United States entered World War I, England was able to sell large quantities of bonds in this country and thus acquired dollars with which direct payment could be made for her imports. After the United States entered the war, our government borrowed from its citizens through the sale of Liberty Bonds and placed part of the proceeds at the disposal of the English government and the governments of other allied countries. During World War II, the government of the United States, using funds acquired by borrowing and taxation, purchased large quantities of foods, materials, and equipment for war and turned these commodities over directly to the governments of England, Russia, China, and other countries under the lend-lease program. In this way, these countries were enabled to obtain vital imports from the United States without making any immediate payment for them. Thus in times of emergency the normal uses of foreign exchange and the ordinary process for the determination of exchange rates may be almost completely suspended.

QUESTIONS AND PROBLEMS

1. Discuss the problem which is involved in the settlement of international obligations.

2. "Most payments between countries, like those within countries, are made by means of credit instruments." Show whether you agree.

3. How would an American exporter obtain payment for a bill of goods sold in England? Explain.

4. "The use of bills of exchange is convenient for both exporters and importers." Explain.

5. "A bank takes a considerable chance in buying a bill of exchange from an exporter." Show whether you agree.

6. How does a bank recover the funds which it pays to an exporter for a bill of exchange? Explain.

7. "If our exports to England greatly exceed our imports from England in a given period of time, a large flow of gold from England to the United States will be sure to occur if both countries are on the gold standard." Do you agree? Explain.

8. "The rate of exchange between two countries is a price and it is determined in much the same fashion as other prices." Show whether you agree.

9. Show the nature and significance of the gold par of exchange between gold-standard countries.

10. "Rates of exchange between countries on the gold standard can vary around the gold par of exchange only within rather narrow limits." Do you agree? Explain.

11. "Fundamentally, exchange rates are determined by demand and supply conditions." Explain.

12. "When the exchange rate between two countries on the gold standard departs from par, two forces operate to bring the rate back toward the par level." Do you agree? Explain.

13. "The influences which tend to bring the rate of exchange between gold-standard countries back toward par operate much less perfectly in practice than in theory." Show whether you agree.

14. Distinguish between the gold par of exchange and the purchasing-power par.

15. "The exchange rate between two countries fluctuates more widely around the purchasing-power par than around the gold par." Do you agree? Explain.

16. "When two trading countries are off the gold standard, there is nothing to bring the exchange rate back toward par once it departs from that level." Show whether you agree.

17. "When governments go in for foreign exchange control, the normal uses of foreign exchange and the ordinary process for the determination of exchange rates are almost completely suspended." Do you agree? Explain.

18. "If American tourists visit foreign countries when prices are high in this country, they register a gain in connection with both the price level and the rate of exchange." Show whether you agree.

19. The rate of exchange between the United States and England is now above par but is expected to go below par in the near future. Would you advise an exporter to sell goods in England now or wait a while? Explain.

REFERENCES FOR FURTHER READING

Bradford, F. A., *Money and Banking*. New York: Longmans, Green & Company, rev. ed., 1941, Chaps. 15, 16.

Furniss, E. S., *Foreign Exchange*. Boston: Houghton Mifflin Company, 1922, Chaps. 1–5.

Peterson, J. M., Cawthorne, D. R., and Lohman, P. H., *Money and Banking*. New York: The Macmillan Company, 1941, Chaps. 29, 30.

Thomas, R. G., *Our Modern Banking and Monetary System*. New York: Prentice-Hall, Inc., 1942, Chaps. 24, 37–41.

Westerfield, R. B., *Money, Credit and Banking*. New York: The Ronald Press Company, 1938, Chaps. 24–26.

XXVII

International Trade

In Chapter V, in discussing the subject of specialization, we called attention to the fact that specialization and the exchange of products occur on the basis of geographical areas, such as communities, states, regions, and countries, as well as among individuals and firms. In the present chapter we examine the subject of geographical specialization and trade in greater detail, and pay particular attention to the specialization and trade which goes on between countries.

THE THEORY OF TRADE

The nature and extent of interregional and international trade, and the reasons for its existence, will be understood most readily if we begin by stating the theory of trade on the basis of a number of assumptions as to the conditions which exist within and between regions and countries, even though these assumed conditions may not always be realized in practice. For the present, then, we assume that the various agents of production are completely mobile between occupations and places within each region or country, but that they are completely immobile between regions or countries. The agents of production are also taken to be perfectly divisible in each region or country, so that there are no advantages or disadvantages associated with the size or scale of particular enterprises, although there will be in each region or country an optimum proportion of the agents of production in each field of production. The prices of all economic goods are supposed to be equal to the long-run minimum average costs of producing them, all agents of production are remunerated at rates which represent the marginal productivity of these agents, and, in general, the various conditions of competition prevail in all lines of production. Finally, costs of transporting goods between regions or countries will be ignored for the most part.

The Basis of Interregional and International Trade. Under the assumed conditions, interregional and international trade depends in the first place on differences between regions or countries with regard to the productive agents which they possess. As we noted in Chapter V, the location of par-

ticular industries in given geographical areas is accounted for by some one or more of such factors as the existence of a suitable supply of power, the availability of raw materials, nearness to markets, the presence of a supply of labor of appropriate grades, the availability of capital funds for investment, favorable climatic conditions, and the advantage of an early start.

Thus some regions or countries have enormous deposits of excellent coal, iron, petroleum, lead, zinc, copper and other important industrial materials, while other regions or countries have very few resources of these kinds. Timber and water power are plentiful in some regions or countries but scarce in others. Some regions or countries have populations which are small in relation to their areas and resources and which contain large proportions of individuals who are trained or skilled in a variety of industrial pursuits; others have relatively large populations which are composed for the most part of unskilled workers. Some regions or countries have the benefit of great aggregations of capital for facilitating the processes of production but others are most backward in regard to capital accumulation. Some regions or countries have produced certain products for many years and have a rich background of experience and tradition, while others are inexperienced newcomers in the same fields of production. Even differences in climatic conditions may have a profound influence on the types of economic goods which may be produced in different areas.

Such differences in their endowments of productive agents between regions or countries lead to differences in the efficiency with which different types of economic goods may be produced from one area to another, and hence to variations in the cost per unit of producing these goods. A region or country which has large and excellent supplies of iron and coal, an adequate amount of capital, and large numbers of able workers of the necessary types would be expected to turn out iron and steel products with greater efficiency and at lower cost than another area which has either small quantities or poor qualities of iron and coal, whether or not the latter area had the capital and labor requirements at hand. Again, a region or country which has large quantities of low-grade labor and little else may be able to turn out very cheaply products requiring much hand labor, while another region or country with better labor and large quantities of capital and land may find its labor so valuable for other purposes that it cannot afford to produce any of these handmade products, and so on. Where significant differences in efficiency and cost of production prevail as the result of differences in endowments of productive factors between regions or countries, opportunities for profitable trade between them exist under the assumed conditions.

Since advantages and disadvantages with respect to the endowments of productive agents are scattered throughout the world in such a way that no region or country can possibly be the best place in the world to produce

all kinds of economic goods, all regions and countries are likely to have an opportunity to gain by trade with other areas. Some imagination is required to visualize conditions under which there would be no reason for the existence of trade between two regions or countries. If demand and supply conditions for all finished goods and agents of production were exactly the same in both areas, no trade would be carried on between them, but this situation could scarcely be found in practice in view of the differences which exist between regions and countries in the matter of their endowments of productive agents. Again, if the difference between two areas with respect to endowments of productive factors were precisely offset by a difference between them in the demands for economic goods, a no-trade condition could exist. However, since differences in the endowment of productive agents between regions or countries are likely to be greater than differences between them with respect to the demand for finished goods, the chances are excellent that opportunities for profitable trade will exist between areas with different kinds, quantities, and qualities of the productive agents.

Absolute Advantages in Trade. Whenever a given country or region, because of its particular endowment of productive agents, is able to produce an economic good with greater efficiency than other areas, it is said to have an absolute advantage in the production of that good. For example, Brazil has land, labor, and climatic conditions which are well suited to the production of coffee, while this commodity could be produced in the United States only with the greatest difficulty, if at all. The United States, on the other hand, has excellent facilities for producing automobiles, but Brazil, in addition to being short of workers of the proper skill and training, and perhaps capital as well, lacks the necessary supplies of iron and coal which the United States has in great abundance. Thus, if we consider only these two countries, the United States has an absolute advantage in the production of automobiles, while Brazil has an absolute advantage in the production of coffee.

Comparative Advantages in Trade. A region or country will not necessarily produce and export all economic goods in the production of which it has an absolute advantage over other areas, nor will an area with no absolute advantages necessarily refrain from exporting. That is, a region or country which has many absolute advantages will probably find that it has a comparative advantage in certain fields of production, and even an area which has no absolute advantages is likely to have a comparative advantage in the production of certain economic goods. Suppose, for example, that both the United States and England have the general types of land, labor, and capital necessary for the production of both wheat and flax, and that the United States has an absolute advantage in the production of both goods. Suppose further that the cultivation of flax requires a great deal of

labor and relatively little land and capital, while the cultivation of wheat requires much land and capital and relatively little labor. Finally, let us assume that labor is relatively scarce and high-priced in the United States as compared with the land and capital, while in England the reverse is true.

In this situation, the United States is likely to have a comparative advantage in the production of wheat, and England will probably have a comparative advantage in the production of flax. In spite of our great physical efficiency in the production of flax, the money cost of raising flax in the United States may well be greater per unit of product than its cost in England, since flax raising requires a good deal of labor which is relatively scarce and high-priced in the United States. On the other hand, the money cost of wheat per bushel is likely to be lower in the United States than in England, since wheat raising requires only relatively small amounts of our high-priced labor. The United States, although it has an absolute advantage in both lines of production, may therefore find it profitable, as between these two commodities, to concentrate on wheat raising and to obtain flax, or the economic goods derived from it, by importation. In opposite fashion, England, although it has an absolute disadvantage in both lines of production, may well go in for flax raising on a large scale while obtaining large amounts of wheat by importation.

The Law of Comparative Costs. Under the conditions which have been assumed, we conclude that a region or country tends to export economic goods whose production requires large quantities of productive agents which are abundant and cheap in that area and small amounts of productive agents which are scarce and dear. Conversely, a region or country tends to import economic goods whose production requires large amounts of productive agents which are scarce and dear in that area and small quantities of productive agents which are abundant and cheap. Not only does it not pay a region or country to produce every type of economic good which can be produced there; it does not even pay a region or country to produce all types of economic goods which it can produce with greater physical efficiency than other areas. A region or country can often gain, in terms of both money costs and the quantities of all types of economic goods which it will have available for consumption, by giving up the production of articles which it can turn out with only slightly greater physical efficiency than other areas in order to use the productive agents, which would otherwise be devoted to such industries, in other fields of production in which it has large advantages in efficiency over other areas.

The formal principle of economics which covers this situation, known as the Law of Comparative Advantage, states that *a region or country tends to export those economic goods in the production of which it has a comparative advantage and to import those economic goods in the production of which it has a comparative disadvantage.* Fundamentally this law means

that the economic principles of specialization and exchange are valid between regions and countries as well as between individuals and firms. The law is not intended to suggest that a region or country should limit itself to producing the one or two economic goods in the production of which its absolute advantage is greatest or its absolute disadvantage is least. However, in order to follow the law consistently, most countries, if not smaller regions, would have to give up some of the lines of production which they are now attempting to carry on. That is, most countries, because of severe governmental interference with the course of international trade, now attempt to produce at home some economic goods in whose production they actually have a comparative disadvantage.

Trade and the Prices of Finished Goods. Under the conditions which have been assumed, one effect of trade between two regions or countries is to equalize prices between them for goods which are able to enter into trade. Suppose that Argentina and England are two countries under discussion and that, in the beginning, they are not trading with each other. Such raw products as beef, hides, wheat, and wool are cheap in Argentina, for this country has the great land expanses and adequate amounts of the other agents necessary for the production of these goods; at the same time, manufactured articles are high in price since Argentina is short of the facilities necessary to produce them. In England just the reverse is true. England has the large aggregation of capital, abundant supply of trained labor, and limited land space which under ordinary conditions incline a country toward industrial or manufacturing production, while the same limited land space places a serious obstacle in the way of agricultural production.

These differences between the countries with respect to the prices of the two classes of goods disappear when unrestricted trade springs up between them. The various agricultural products pass from Argentina, where they are relatively plentiful and cheap, to England, where they are relatively scarce and dear; at the same time, various manufactured products move from England, where they are relatively abundant and low in price, to Argentina, where they are relatively limited in amount and high in price. Under the assumed conditions there is nothing to stop the flow of goods until each class of products sells for the same set of prices in the one country as in the other. If we bring into the picture the costs of transportation and other obstacles to trade, the equalization of the prices of the goods between the countries is not complete. Rather, wheat, let us say, moves to England until its price there differs from the price in Argentina by not more than the cost of transporting a unit of wheat between the countries, or until its price in England is higher than its price in Argentina by not more than the cost of transportation plus the amount of any import tax which England imposes on wheat, and so on.

Trade and the Prices of Productive Agents. There is also a tendency toward the equalization of the prices of the different varieties and grades of the productive agents in two regions or countries when they trade with each other. Exporting agricultural products to England, for example, increases the demand for the productive agents which are relatively abundant and cheap in Argentina and raises their prices, while the dependence of England on imports of agricultural products from Argentina relieves the pressure on those productive agents which are relatively scarce in England and lowers their price. In similar fashion, exporting manufactured goods to Argentina increases the demand for those agents of production which are relatively plentiful in England and raises their prices, while the prices of these same agents in Argentina tend to fall as imports of manufactured goods from England brings about a decrease in the demand for the agents in Argentina. Thus, with agents necessary to agricultural production rising in price in Argentina and falling in price in England and with agents necessary to industrial production rising in price in England and falling in price in Argentina, the tendency toward the equalization of the prices of the agents of production in the two countries is clear.

The Gains from Specialization and Trade. From the point of view of maximizing production it would be most unfortunate if no trade existed between regions or countries whose endowments of the productive factors differed considerably. In each region or country, waste and loss would attend all attempts to produce economic goods which it was indifferently suited to produce. On the other hand, if geographical specialization and the equalization of the prices of economic goods could be complete between regions or countries, all endowments of productive agents would be used for the most appropriate purposes, and real income and scales of living would be at a maximum in the trading areas. Each region or country would gain by receiving from other areas economic goods which could not be produced at all at home, or which, although they could be produced at home and perhaps with satisfactory physical efficiency, could be obtained more cheaply from the other areas. The effect of trade would be to remove the disadvantages which were originally imposed on the regions or countries by the uneven distribution of productive resources among them. In practice, of course, interregional and international trade is not fully developed and this effect is only partly achieved.

When two regions or countries engage in trade, both of them tend to gain, but they do not necessarily gain to the same extent from the interchange of goods. The extent of the gain realized by each region or country depends on what has been called the play of reciprocal demand; or, in other words, the strength of the demand of the one area for the economic goods which it imports as compared with the strength of the demand of the other area for the economic goods which the first area exports. For example,

if Country A needs its imports from Country B very badly and cannot get along without them, while Country B does not care much about the goods obtained from Country A or could just as well get them from some other country, the terms of trade will be relatively unfavorable to Country A. Both countries will gain from the trade, but the lion's share of the gain will go to Country B. Again, any increase in the demand of one country for the goods of another country, other things remaining the same, will tend to alter the terms of trade in a manner unfavorable to the first country. That is, the first country will have to give up relatively more of the goods which it exports in order to get the goods which it desires to import.

Assumptions concerning Mobility and Conditions of Equilibrium. It is now time to relax some of the assumptions under which the theory of interregional and international trade has been stated. In the first place, agents or production are not completely mobile within regions or countries, nor are they completely immobile between regions or countries. Many agents of production achieve at least a fair degree of mobility between places and occupations over long periods of time, but they are notoriously immobile in short periods of time. Even in the long run some agents of production may be more mobile between regions or countries than others are within these areas. Thus, for example, capital funds may move about more readily between regions or countries than some grades of labor move within a particular region or country. Even if the other conditions of competition prevail within a region or country, lack of mobility of the productive agents will prevent the existence of equilibrium conditions with respect to the prices of finished products and those of the agents of production.

In other words, the prices of finished products do not always equal their average costs of production per unit under competition, nor do the rates at which the agents of production are remunerated always equal the marginal productivity of the agents. If excessive quantities of productive agents are tied up more or less permanently in certain fields of production, some industries may be unable to operate their fixed productive facilities at any rate which will give them a price adequate to cover average cost of production per unit of product, and they may be willing, because of heavy fixed costs, to produce and sell their goods at a price lower than average cost per unit. If this were true, articles which could not be exported at a price high enough to cover average cost of production per unit may enter into trade at such prices below cost. In somewhat similar fashion, if a certain occupation or type of work is overcrowded, the workers are unable to move out of it, and the wage which is paid is lower, in a certain period of time, than the marginal productivity of the employed workers, the price of the good produced by the workers may be artificially low. Because the workers are underpaid it may be possible to export the good to other areas, although

this result would not occur if the workers were paid at the rate which would prevail under equilibrium conditions.

Assumptions concerning Competition. If many products are produced under conditions of monopoly or monopolistic competition, certain effects on interregional and international trade may be expected. We concluded in Chapter XVI that, if the firms in an industry are many and small so that the entrance of new firms or the exit of old firms is relatively easy, the price of an economic good under conditions of monopolistic competition may only equal average cost of production per unit in the long run. However, this cost of production and the volume of output achieved may be quite different from what they would be under full competitive conditions. And, of course, the price charged by a monopolist, duopolist, or oligopolist may also differ considerably from that which would prevail under competition. We must conclude that a good which might be exported from an area if the industry producing it were organized competitively may not be exported if the industry operates under noncompetitive conditions. On the other hand, in a shorter period of time in which an industry has a fixed amount of plant and equipment, a monopolist or monopolistic competitor may sell a certain volume of products abroad at prices little more than enough to cover variable costs, in order to obtain the advantages to be derived from operating the fixed plant and equipment at full capacity, while selling only a limited quantity of the products at home at highly profitable prices.

We have also seen, in Chapters XVII–XX, that industries organized as monopolies or under other noncompetitive conditions will take smaller quantities of the productive agents at given prices than the same industries would be expected to take if they had the same size under competitive conditions. This is because a monopolist or a monopolistic competitor can afford to go on hiring units of a productive agent only as long as the goods turned out by each additional unit of the agent will contribute marginal revenue in excess of their marginal cost, while competitive producers can go on hiring units of a productive agent as long as its cost per unit does not exceed its marginal physical product multiplied by the price of the product. This conclusion depends in turn upon the fact that the price which a monopolist or a monopolistic competitor can obtain for his product varies inversely with the amount of the good which is produced and offered for sale, while the price which a competitive producer can obtain for his product is independent of his volume of output because of his insignificant size in relation to the industry as a whole.

When monopolists and other noncompetitive producers limit artificially the quantities of the agents of production which they employ, large quantities of the agents are left for use in other industries, the price at which the entire quantity of each productive agent can find employment is lower

than it would otherwise be, costs of production in various industries are affected, and the possibilities of importing and exporting various economic goods are affected in turn. In short, the production and sale of goods under monopoly and other noncompetitive conditions are likely to affect the content of an area's external trade whenever they affect the prices at which finished goods are sold or the rates at which the agents of production are remunerated.

Assumptions concerning Divisibility of Productive Agents. We have also assumed the existence of perfect divisibility of the productive agents, so that a large-sized combination of the productive agents would have no advantages over a small-sized combination of these agents in any industry. In practice there are advantages for large-scale concerns over small-scale concerns in many types of productive activity, and these advantages depend upon the lack of divisibility of the productive agents. That is, large units of certain agents may be more efficient in relation to their cost than small units, while the smallest units of certain agents that can be operated may call for rather large size on the part of the concern as a whole.

The advantages of large-scale production make possible another phase of specialization between regions or countries, since large-scale concerns in the various industries do not require exactly the same combination of the productive agents. Moreover, the development of different types of large-scale production in different regions or countries alters the demand for and relative prices of the various productive agents in these areas. The result is that, even if trade would not exist between certain areas when all their industries were organized in small-scale units, it may spring up between them as many industries become large-scale and some products of large-scale industry become cheaper than before in one area while other products of large-scale industry become cheaper than formerly in the other area. In such cases trade removes to some extent the disadvantages under which the regions or countries would otherwise labor because of the lack of divisibility of the productive agents.

THE BALANCE OF TRADE

The Balance of Imports and Exports. Another fundamental principle in connection with interregional and international trade is found in the fact that the imports of a region or country must equal its exports over any considerable period of time. Ignorance concerning this principle seems to be widespread, as the many campaigns carried on by communities and states to induce people to "buy at home" will testify. However, the existence of the principle as between countries comes as a particular surprise to people who are accustomed to reading in the newspapers that the United States has an excess of exports over imports in each year and enjoys a so-called favorable balance of trade. The trade statistics which appear in the newspapers ordi-

narily refer only to the merchandise or commodity imports and exports of the country, which pass through the customhouses and are readily counted and measured. Thus if one is willing to disregard everything save the merchandise imports and exports of a country, it is quite possible for the country to have an export balance of trade not only in a given year but year after year.

Invisible Imports and Exports. However, the trade of a country is not likely to consist entirely of merchandise items. Most countries have a considerable volume of service imports and exports, and our proposition concerning the equality of imports and exports has to do with trade in both commodities and services. The service items in international trade are often referred to as invisible imports and exports, since it is frequently impossible to see the item which is imported or exported. If mistakes are to be avoided, we must be careful in classifying the invisible items of trade as imports or exports. Suppose, for example, that people in the United States send funds to relatives or friends in Europe during the coming winter. How should this movement of funds be classified? Most students would say at once that the item in question is an export, since all that can be seen is the movement of funds out of this country. However, the item is really an import.

This startling conclusion is reached in the following manner. When the United States exports merchandise, commodities leave this country and payments for them are made to this country. Therefore, in dealing with the invisible items of trade, we classify as exports only those items which involve payments which move to this country. When the United States imports merchandise, commodities come to this country and payments for them move to other countries. Therefore, in classifying the invisible items of trade, we consider as imports any items which involve payments moving from this country to other countries. These principles are quite simple to apply in practice. When we make a charitable contribution of funds to the people of another country, a payment of funds moves from this country to the other country, and the item is an import. Just what it is that we import cannot be seen. Perhaps, though a cynic may doubt it, it is the love and affection of the people of the other country. At any rate, it should be their better health and physical condition. Conversely, if another country sent funds to this country to help in feeding starving Americans, the item would be an export from our point of view.

If we pay other countries freight charges for carrying our goods over the oceans in their vessels, the items are invisible imports. If similar payments are made to us, they are invisible exports. When immigrant boys come to this country, make good, and send funds back to relatives in the old countries, the United States has invisible imports. Similar remittances to this country are invisible exports for us. If our tourists visit France and

we send funds to them to finance their travels, we import the French scenery, hotel service, guide service, taxi service, or what not, because the items are invisible imports for us. Similar payments to this country to finance the travel of foreigners here are invisible exports from our point of view. If the Bank of England withdraws a deposit from the Federal Reserve Bank of New York and has the funds sent to England, the United States has an invisible import, while the opposite transactions would be an invisible export for the United States, and so on down the long list of items which are sometimes identified as invisible items of international trade.

The Significance of Invisible Imports and Exports. Invisible imports and exports are important because trade in them often goes far toward canceling the results of trade in merchandise items. A country which has an import balance of trade with respect to merchandise is likely to have an export balance of trade with respect to the invisible items, and vice versa. In the case of the United States, our export surplus with respect to commodities has usually been offset to a considerable extent by our import surplus with respect to the invisible items. It should not be difficult to see why this is true. The people of the United States quite often make charitable contributions to other countries, but these contributions are never made to the United States by other countries. In ordinary years the number of Americans traveling in foreign countries and requiring funds is much greater than the number of foreigners traveling in the United States. Many immigrants have come to the United States in the past and have later desired to send funds to persons in other countries, while comparatively few American boys emigrate to other countries to make good and send funds back home. Other countries have had merchant marines which were much better developed than that of the United States so that the total of freight charges which we have had to pay to other countries ordinarily has exceeded by a considerable sum the total freight charges which other countries have had to pay to us in a given year, and so on.

Correcting an Excess of Imports or Exports. But, it may be asked, may not a country have a surplus of exports or imports in a given year with respect to both merchandise and invisible items? The answer to this question is in the affirmative, but a country cannot keep on having a surplus of imports or exports over a period of years. If the United States in a given period exports more of both commodities and services than it imports, as a result of this trade other countries owe us a net balance. If the situation continues for a while, how will such net balances be paid? If the United States and other trading countries are on the gold standard, they will be paid, sooner or later, in gold. They may be paid later rather than sooner if the United States is willing to lend heavily to the countries which are purchasing our exports, but interest and principal are supposed to be paid to this country later on in connection with such loans, and the necessity of

gold movements is only postponed and not eliminated. When these gold movements occur, a train of events now familiar to us is set in motion.

As gold moves to the United States, the quantity of money and credit in circulation in this country increases while that of other countries decreases. Our price level rises and their levels fall. The United States becomes a good place in which to sell and a poor place in which to buy. The reverse is true of the other countries. The exports of the United States decrease and its imports increase, while again the reverse is true elsewhere. As our imports increase and our exports decrease, our former surplus of exports with respect to commodities and services is corrected. However, all the qualifications which were noted in the preceding chapter should be borne in mind in connection with this whole process.

If the United States and other trading countries are off the gold standard, but the United States still has an export surplus with respect to commodities and services taken together, gold is not likely to come to this country in settlement of trade balances. If it does not, rates of exchange on other countries will tend to fall so low in this country that our exports will be discouraged, our imports will be stimulated, and once more the surplus of exports will be eliminated sooner or later. Of course, a surplus of imports with respect to commodities and services taken together tends to be corrected by the same mechanisms working in the opposite direction.

In conclusion we see that a surplus of exports or imports with respect to commodities alone can be maintained by a country indefinitely but that such a surplus means very little. An export or import surplus in connection with both commodities and services may exist for a country in a particular year, but over a period of a few years such a surplus tends to be eliminated. Finally, if we include in our discussion of trade balances such items as gold movements, security movements (purchases and sales) between countries, and short-term borrowings and lendings between countries, the trade of a country, or at least its transactions calling for payments into and out of the country, is in balance in each individual year. This will be seen in the case of the United States a little later in the present chapter.

OBSTACLES TO TRADE

In view of the fact that regions and countries differ considerably with respect to their endowments of the productive agents and that production and scales of living the world over could apparently be maximized through geographical specialization and exchange, we should expect that the various regions and countries would seek to promote the development and expansion of trade. As a matter of fact, however, most countries and many smaller regions have sought to limit and restrain their trade with other areas. They have usually been interested in restraining imports rather than exports, but we know from our discussion of the balance of trade that exports will also

have to remain at a low level in the long run if imports are severely restricted.

The Protective Tariff. Historically, the most important device for restricting international trade has been the protective tariff, under which various articles produced abroad are made dutiable at rates designed to exclude these products from the country, so that the domestic market is left free for exploitation by domestic enterprises turning out the same types of goods. For example, if a certain article can be obtained from England at a cost of 75 cents per unit (including transportation costs) but cannot be produced in this country for less than $1.00, then a duty of, say, 50 cents per unit is levied on this article, so that our wants for the good will be satisfied by domestic production, if satisfied at all. When similar duties of varying amounts are applied to hundreds or thousands of articles, the result is a protective tariff. While some goods may continue to be imported despite the duties which are applied to them, the purpose of the protective tariff is clearly the exclusion of foreign goods.

Other Restraints on International Trade. Sometimes the same result is achieved by paying bounties or subsidies to domestic producers, which make it possible for them to sell their wares at prices below full costs of production. Such low prices for the domestic goods make it difficult for foreign producers to compete in the domestic market, and relatively easy for the subsidized producers to sell in foreign markets, if foreign countries do not take steps to exclude the goods. It is also possible to protect domestic producers by means of sanitary regulations applied to imported food products, veterinary laws, regulations requiring that imported goods bear a distinctive mark indicating their country of origin, and regulations governing the granting of public contracts, which require that domestic materials must be used or make it necessary for materials to come up to certain specifications which foreign materials can hardly meet. In modern times, international trade is often restricted by means of import quotas, or laws and decrees which limit (by value, quantity, or weight) the imports of certain articles that may enter the country within a specified period of time. Imports or exports of various articles may also be forbidden unless the government grants permits or licenses covering the specific transactions.

The government of a country may control imports and exports by controlling transactions in foreign exchange. Under one common type of foreign-exchange control, the government requires all exporters to sell all bills of exchange drawn on their foreign customers to a governmental agency, which in turn doles out foreign exchange to importers for the importation of approved goods. In other cases, clearing agreements are used to control the volume and content of international trade. A fixed exchange rate between the monies of two countries is agreed upon, and a governmental agency is set up in each country to handle trade transactions. Im-

porters then pay the domestic authority of their country for goods imported, while exporters receive payment from the domestic authority and not from foreign importers. In this way the obligations arising from trade are settled within each country, and the use of foreign exchange is avoided as long as the total imports and exports of the countries are kept in balance. It is also clear that the governments engaging in such agreements are able to control the volume and content of their foreign trade and afford protection to domestic industries.

Internal Restraints on Trade. Even within countries, interferences with trade are fairly common. We are often asked to patronize the home merchants of our city instead of going or sending out of town for goods, and we are exhorted to buy the products of our own state in preference to those from other states. Moreover, in the great depression after 1929, many of our states placed direct restraints on interstate trade. The devices used included scores of kinds of discriminatory taxes on the products, firms, and commercial motor vehicles of other states; sanitary, quarantine, and inspection laws; grade, standard, and label requirements; restrictions on the movement of workers; requirements excluding all but legal residents of the state from public payrolls; laws requiring the expenditure of public funds within the state; and regulations limiting or forbidding the exportation of products of state natural resources.

The "Buy at Home" Fallacy. In view of the results which widespread interregional and international trade might be expected to produce, why do regions and countries so commonly restrict and limit their trade with other areas? All too often the policy of restrictionism is based on the "buy at home" fallacy, or the erroneous idea that great economic advantages may be derived from having the inhabitants of a given region or country do as much as possible of their buying within the home market, instead of outside. The supposed advantages of buying at home are many and varied, but we can consider only a few of them here. It is argued that the wealth of the home area can be built up in this way, that the home market can be expanded in size, that the home area can have more industries and hence provide more employment for workers, and that wages can be raised or at least kept from falling to levels prevailing elsewhere. It will be easy for us to see that these "advantages" are imaginary, rather than real, if we concentrate our attention on the goods aspect of trade and disregard monetary considerations.

In the first place, it should be remembered that the wealth of a country is not composed of its money but exists in the form of a stock of scarce, useful, material, and transferable things. If we really wish to build up the wealth of the home area, the way to do it is to spend our money incomes in such a way that we maximize the quantity of these scarce, useful, material, and transferable things in our possession. This means spending our

money at home if we can get more articles of wealth for it there than elsewhere, or even as many, but it also means buying in other regions and countries if we can get more for our money in that way than by buying at home. From the point of view of increasing our real wealth, we cannot do better than to exchange our money incomes for the largest possible quantities of other things which we desire.

The arguments to the effect that we can expand our markets, and increase industries, employment, and wages by buying at home, would be more convincing if we could refrain from buying in other regions and countries and still retain our ability to sell goods in these other areas. However, we know that this cannot be done. If the people of a given region or country will not import, they cannot continue to export. If they buy almost exclusively in the home market, they must sell almost exclusively in the home market. Thus, through buying at home, we build up our markets at home by losing our markets in other areas, we gain industries which produce for the domestic market by losing or contracting industries which produce for outside markets, we increase employment in the former industries by decreasing employment in the latter industries, and so on.

The net effect of buying almost everything at home is to make it necessary to produce almost everything at home and, since advantages in production are widely scattered over the regions and countries of the world, this means that some goods will be produced efficiently at home and some inefficiently. It is difficult to see how, by maintaining a combination of efficient and inefficient industries at home, we can build up the wealth of the home area, or increase its industrial activity, employment, and wages, in a manner which would compare favorably with the results achieved under another system in which we would produce at home only those things which could be turned out most efficiently there and would obtain other things by trade from other areas in which they too could be most efficiently produced.

We should also note that employment is a means to an end, rather than an end in itself, and that precisely the same thing is true of money wages. Regardless of how many pieces of money our workers receive, it is impossible to believe that their real wages can be higher when they are compelled to spend their money wages for a combination of efficiently and inefficiently produced goods than when the money wages are spent entirely for efficiently produced goods obtained in part through domestic production and in part by exchange with other areas.

Other Bases of Restrictionism. Sometimes the policy of restricting trade rests on bases other than those which have just been examined. In some cases, imports have been restricted in order to protect "infant industries" in the home area. If the young and temporarily inefficient industries are protected from the competition of lower-priced products from outside areas

for a time, they may develop in size and efficiency until their products can be turned out more advantageously than those of similar industries in other regions or countries. It is anticipated that the gains to be realized after the infant industries have grown up will be more than sufficient to compensate for the losses sustained while their protection is necessary. Thus, from the long-run point of view, the region or country will really be following the policy of producing at home only those things which can be produced most efficiently there. This argument for restricting trade in order to protect infant industries would be entirely valid except for two circumstances. In the first place, it is practically impossible to decide in advance just which of an area's infant industries will eventually become strong and efficient. Second, it is often true that industries, no matter how great and efficient they may become, do not feel that they are sufficiently grown up to surrender protection and face foreign competition.

In modern times the policy of restrictionism, as applied to international trade, has often rested on the desire for national preparedness, self-sufficiency, and independence in times of war. Under a system of free international trade, each nation would be highly dependent upon other nations for certain commodities, and would be subject to the ever-present danger of having her supplies of these commodities cut off in wartime. Accordingly it is argued that we should exclude most foreign products and thus ensure within the country the production of as many essential products as possible. What if this policy does mean that, to a certain extent, our productive resources will be used ineffectively and a sum total of commodities smaller than the maximum will be produced, just so long as our national security is promoted? This argument is widely accepted but, in addition to admitting the direct uneconomic effects of restricting international trade, it also overlooks the possibility that, if nations traded freely and were highly dependent upon one another, the likelihood of war might be lessened to a marked degree. In conclusion we should note that, whatever may be the arguments on which the policy of restricting trade is based, the result of the policy is a failure to maximize production and scales of living. Whether this result should be considered "good" or "bad" must be decided on other and nonscientific grounds.

International Trade and Economic Systems. This may be a good time to remark that the principles of international trade are of general application to countries regardless of whether their economic systems are capitalist, socialist, communist, or fascist. No one country can be made the best place in the world in which to produce all kinds of economic goods merely by changing from democracy to dictatorship or from capitalism to one of the other isms. Under any system a country which will not import cannot export. The Law of Comparative Advantage will operate for any system in the absence of governmental decisions to the contrary. Any country, regard-

less of the type of economic system which it has, may decide to restrict its foreign trade and to develop as high a degree of economic self-sufficiency as possible. And finally, when any country decides to follow a policy of restrictionism, the results will be those which we have specified, without regard for the type of economic system which exists in the country.

THE INTERNATIONAL TRADE OF THE UNITED STATES

The Balance of Payments in 1940. In turning to a brief discussion of the international trade of the United States, we present, in Table 22, an estimate of our balance of international payments in 1940 [1] and note that our exports

Table 22: Balance of International Payments of the United States in 1940*
(*In millions of dollars*)

Item	Exports	Imports	Net Exports	Net Imports
1. Merchandise	$4,021	$2,625	$1,396	
2. Service items				
Freight and shipping	223	327		$ 104
Travel expenditures	81	223		142
Personal remittances	30	120		90
Institutional contributions	—	55		55
Interest and dividends	525	215	310	
Government transactions	28	122		94
Miscellaneous services	164	66	98	
3. Capital movements				
Long-term investments (net)	—	53		53
Flow of banking funds (net)	873	—	873	
Advance payments by British government	720	—	720	
Miscellaneous capital items	—	170		170
4. Gold, silver, and paper currency				
Gold	5	4,749		4,744
Gold earmarking operations (net)	645	—	645	
Silver	4	59		55
Paper currency (net)	33	—	33	
5. Residual items (net)	1,432	—	1,432	
	$8,784	$8,784		

* Source: United States Department of Commerce.

of merchandise amounted to $4,021,000,000, while our merchandise imports were valued at $2,625,000,000. Hence we had an export surplus of $1,396,000,000 with respect to merchandise. In examining the service items we should recall that payments into the United States are classed as exports, and payments out of the United States are classed as imports. Thus the figures for freight and shipping service indicate that we paid other countries $327,-

[1] After 1940 the international trade of the United States was seriously disrupted by World War II, and the government withheld most statistical data concerning this trade for reasons of national security.

000,000 for these services in 1940 and received only $223,000,000 for this item. In similar fashion, the data for travel expenditures show that we paid $223,000,000 to other countries and received $81,000,000 from other countries. Our net "imports" on account of travel expenditures were thus $142,000,000. The interest and dividends item indicates that we had a net "export" of $310,000,000, since other countries paid us $525,000,000 while we paid them $215,000,000.

With regard to all the service items, the United States had "imports" in 1940 which exceeded her exports by $77,000,000. This means that our payments to other countries on account of these items exceeded their payments to us by that sum, and our surplus of merchandise exports, requiring net payments to us, was canceled to this extent. In ordinary years our excess of invisible imports has come much closer to offsetting our excess of merchandise exports than in 1940, but in that year our trade was brought back to balance largely by a heavy gold movement, and movement of capital funds, toward the United States. Gold movements are classified in the balance of payments as if they were commodity imports and exports; hence the table indicates that the United States received $4,749,000,000 in gold from other countries in 1940 and exported only $5,000,000, although the net gold imports of $4,744,000,000 were offset to the extent of $645,000,000 by earmarking operations.[2]

Movements of capital funds are put into the balance of payments as invisible items. The item of $53,000,000 for long-term investments indicates a net payment of that amount to the people of other countries as the result of purchases and sales of American and foreign securities in 1940, and miscellaneous capital items resulted in a similar payment of $170,000,000. On the other hand, the sum of $873,000,000 flowed to the United States as the result of changes in the foreign deposits and short-term loans of our banks and in the American deposits and short-term loans of foreign banks, while a payment of $720,000,000 was received by the United States from the British government as advance payments for merchandise and as capital assistance for certain American firms which needed new facilities in order to produce for England. Capital movements altogether resulted in a net export, or payment into the United States, of $1,370,000,000. Over the whole balance of payments, we see that merchandise, service items, capital movements, gold, silver, and currency movements, and residual items resulted in payments to the United States of $8,784,000,000 and payments of the same amount from the United States.

Trends in the International Trade of the United States. The composition of the merchandise exports and imports of the United States has been

[2] Earmarking operations involve gold specifically set aside in a country and marked to indicate that it belongs to foreign owners, who may dispose of it at will.

changing gradually over the years.[3] About a century ago, in 1840, the merchandise exports of the United States consisted of about 72 per cent crude materials and crude foodstuffs and only 28 per cent manufactured foodstuffs, finished manufactures, and semimanufactures. Crude materials and crude foodstuffs still made up about 60 per cent of our merchandise exports in the five-year period of 1871–1875, and 45 per cent in the period 1896–1900. As the industrialization of the United States became more and more complete, exports of crude materials and crude foodstuffs continued to decline in relative importance until they amounted to only 31 per cent of total merchandise exports in the period 1926–1930 and to about 21 per cent in 1939. In 1940, under the special conditions created by war, they fell further to less than 14 per cent of the total, while our exports of manufactured and semimanufactured goods increased accordingly.

The trend in the merchandise imports of the United States has been in exactly the opposite direction. Crude materials and crude foodstuffs accounted for only 28 per cent of these imports in 1840, while manufactured foodstuffs, finished manufactures, and semimanufactures accounted for 72 per cent. Crude materials and crude foodstuffs made up 30 per cent of our merchandise imports in the period 1871–1875, 44 per cent in the period 1896–1900, over 49 per cent in the period 1926–1930, and about 51 per cent in 1940, while our imports of manufactured goods and semimanufactures have declined accordingly.[4]

The relationship between the total exports and imports of merchandise has also been changing for the United States over the last century and a half. For many years the United States was what is sometimes called an immature debtor nation, which means that the United States was borrowing large quantities of capital funds from other countries and that the new capital funds coming to this country in each year tended to exceed the payments which we had to make to other countries on account of the interest and principal on old loans. Such net inflows of capital funds affected our annual balance of international payments as export items, since they involved payments to this country. In order to achieve a balance annually, our total imports had to exceed our other export items by the amount of the net inflow of capital funds in each year. Hence it is not surprising that the merchandise imports of the United States exceeded the merchandise exports in most years during this period. In the eighty-six years from 1790 to 1875 there were only fourteen years in which our merchandise imports failed to exceed our merchandise exports.

[3] Data with regard to the individual commodities which are leading imports and exports of the United States, and with regard to the countries which are leading markets for our exports and sources of our imports, are not presented because it is thought that they would be of little value for students of economic principles.

[4] These figures on the changing composition of the merchandise imports and exports of the United States are from the *Statistical Abstract of the United States*, 1941, p. 533.

Eventually the United States became a mature debtor nation, and it had to make annual payments to other countries on account of interest and principal on old loans which exceeded the amounts of capital funds coming to this country in connection with new loans. Since these net outgoing payments ranked as imports in the balance of international payments of the United States, our exports had to exceed our other imports annually by the amount of these payments. In the thirty-nine years from 1876 through 1914, the merchandise exports of the United States exceeded the merchandise imports in all but three years. During World War I, the United States not only got out of debt to other countries but became a creditor nation on her own account—albeit an immature creditor nation, for our new loans of funds to other countries in each year tended to exceed the payments which other countries had to make to us on account of the interest and principal on old loans. Thus, as an immature creditor nation, we had the same annual net outflow of capital payments which marked our years as a mature debtor nation. We continued to have an export surplus with respect to other items in the balance of payments, and our merchandise exports exceeded our merchandise imports by a considerable amount in each year from 1915 through 1929.

In 1929 the merchandise exports of the United States amounted to $5,240,995,000 and the merchandise imports to $4,399,361,000. By 1932 our merchandise exports had fallen to $1,611,016,000 and our merchandise imports to $1,322,774,000, a decline of 69 per cent (by value) for exports and 70 per cent for imports. In the years between 1932 and 1939 the international trade of the United States recovered some of its lost ground, and in 1939 our merchandise exports and imports amounted to $3,177,344,000 and $2,318,258,000, respectively, or 61 and 53 per cent, respectively, of their 1929 levels. In 1940, under the influence of war, our merchandise exports increased to $4,021,146,000 and our merchandise imports to $2,625,379,000.[5]

Many factors helped to account for the lowly status of this country's imports and exports of merchandise between 1932 and 1939. Relatively depressed business conditions here and elsewhere played a large part in the situation. The Smoot-Hawley Tariff Act of 1930 raised considerably the average level of duties applied to articles coming into the United States from other countries, and made it difficult for us to import and, indirectly, to export. Many of the exports of the United States through 1929 were financed by loans to other countries, but these loans were almost completely discontinued after that year. With most countries off the gold standard, unstable foreign exchange rates added to the risks involved in importing and exporting and served to discourage international trade to some extent. Some of the measures which were adopted in the United States after 1932

[5] These data for recent imports and exports of the United States are from the *Statistical Abstract of the United States*, 1941, p. 523.

for the purpose of achieving economic reform or promoting recovery from the depression had an effect on the international trade of this country. Whether or not such measures as the National Industrial Recovery Act and Agricultural Adjustment Act made for recovery from the great depression, they apparently resulted in restricted production and high prices which made it difficult for us to export the usual quantities of industrial and farm products in competition with other countries.

Our tariff policy, as embodied in the Smoot-Hawley Act, led to a host of retaliatory tariff acts on the part of other countries. When the United States was a debtor nation and its exports consisted largely of vital raw materials and foods which other countries had to have, we could follow a high tariff policy with impunity, for other countries were not likely to discriminate against our trade. Now that the United States had become a creditor nation and its exports consisted largely of manufactured goods, many of which our customer nations could produce for themselves at only a relatively slight disadvantage in cost or could obtain from other industrial nations almost as cheaply as from the United States, our high tariff, which placed obstacles in the path of debtor nations which wished to sell goods in the United States, was more likely to lead to retaliatory actions on the part of other countries. Moreover, some countries, acting in the spirit of intense nationalism, tried to develop programs of economic self-sufficiency in the 1930's and restricted imports by means of tariffs, quotas, foreign exchange regulations, and other restrictive devices. Such developments limited our ability to export and, indirectly, to import.

Wartime Controls over International Trade. After 1934, by entering into the so-called reciprocal trade agreements with other countries, the United States made a serious effort to undo some of the mischief which resulted from the Tariff Act of 1930. After five years, however, World War II broke out and the United States became involved in the conflict in late 1941. Wartime conditions brought great increases in the extent to which the international trade of the United States was controlled by the federal government. Even before the United States entered the war, our government set out to accumulate stock piles of certain strategic and critical materials, and entered into agreements with various Latin-American countries for the purchase of all their available supplies of such materials. Moreover, after December 27, 1941, the government assumed complete control over the imports of a number of materials, and these could be imported only by some governmental agency.

Our foreign trade was affected also by the "freezing" of foreign assets. The freezing process was based on an Executive Order administered by the federal reserve banks and the Treasury Department, and it prohibited all transactions within the jurisdiction of the United States in which the country (or its nationals), to which the order applied, had any interest,

after a stipulated date. The freezing process, first used in April, 1940, was extended to country after country as German conquests continued, to Germany and Italy themselves in June, 1941, and to Japan in July, 1941. Naturally, all imports and exports between the United States and any country whose assets had been frozen were automatically prohibited, unless our government saw fit to issue licenses for specific transactions. This was true also of transactions between countries with frozen assets and any third country, if the transactions were to be financed by means of foreign credits held in the United States.

The freezing orders as such were not applied to the Latin-American countries, but something of the same effect was produced by the promulgation in July, 1941, of the Proclaimed List of Certain Blocked Nationals. This list contained the names of persons and firms believed to be nationals of or sympathizers with the Axis countries and located in countries of the Western Hemisphere. Our government forbade all business and financial transactions between citizens and residents of the United States and listed persons and firms, unless specifically permitted by licenses issued by the Treasury Department. The United States had considerable cooperation from Latin-American countries in carrying out this policy, which was aimed quite definitely at depriving the Axis powers of any economic advantages they previously enjoyed from enterprises, investments, and business connections in the Latin-American countries.

The exports of the United States were also subjected to direct control during the war period. In July, 1940, the National Defense Act provided for a general system of export control by means of licenses. This export control system was originally intended to apply to essential raw materials, machine tools, certain chemicals, arms, ammunition, and war goods in general. However, the list was increased rapidly, and soon scarcely anything included in our normal list of exports could be exported without a federal license. The export control system prevented other countries from buying here raw materials and goods which were needed in our war program, but permitted us to send all kinds of goods to countries of the Western Hemisphere which were collaborating with the United States in her war program. Even before our entry into the war, export control enabled us to interfere with and hamper the war activities of the Axis nations.

Finally, the lend-lease policy of the United States had an important effect on our trade. This policy was provided for in the Act to Promote the Defense of the United States, passed in March, 1941. This act authorized the president to sell, transfer title to, lease, lend, or otherwise dispose of various defense goods to other countries whose defense was deemed vital to the defense of the United States. Defense articles in this connection included (1) weapons and munitions of war, (2) machinery, facilities, tools, materials, parts, and supplies necessary to the production, mainte-

nance, and repair of war weapons and goods, and (3) any agricultural, industrial, or other commodity or article for defense. The act specified that lend-lease aid could be given under any terms and conditions which were satisfactory to the president, and that the resulting benefit to the United States could be payment in kind or property, or any other direct or indirect benefit which the president deemed satisfactory. From 1941 through October, 1945, lend-lease aid to Great Britain, Russia, China, and other allied nations amounted to about 38 billion dollars.[6] This outlay represented some 11.7 per cent of American spending for war over the same period.

International Trade of the United States in Wartime. Clearly, the various governmental controls which we have been describing did not all make for a decline in the total volume of our international trade. Some policies, such as the lend-lease program and the over-all purchasing agreements for strategic and critical materials, tended to increase trade. As a result, the total volume of the international trade of the United States, in terms of merchandise, increased year by year from 1939 through 1944, as shown by the data in Table 23. The expanding total volume of trade in this period was largely the result of sharply increasing exports, and lend-

Table 23: International Trade of the United States, 1939–1945*
(*In millions of dollars*)

Year	Exports	Imports	Total Trade
1939	$ 3,177	$2,318	$ 5,495
1940	4,021	2,625	6,646
1941	5,147	3,345	8,492
1942	8,035	2,745	10,780
1943	12,975	5,375	18,350
1944	14,568	6,926	21,494
1945	11,161	6,073	17,234

* Source: *Survey of Current Business*, February, 1945, p. 19, and February, 1946, pp. 25–26.

lease exports of war goods, foods, and industrial materials played a very important part in increasing total exports. In 1944, for example, total exports of $14,568,000,000 included only $3,528,000,000 of ordinary "cash" exports, while in 1945 cash exports amounted to $5,108,000,000 out of total exports of $11,161,000,000.

There were also important changes in the direction of our international trade during the war period. Our exports, consisting largely of lend-lease goods, went mainly to our allies and other countries actively collaborating with the United States. On the import side, too, we were cut off from trade with enemy countries and other countries dominated by them. In particular,

[6] *Survey of Current Business*, February, 1946, p. 24.

we were cut off from our usual sources of supply of such important imports as rubber, tin, and silk in the Far East. Our imports from Europe and Asia, which made up 57 per cent of our total merchandise imports in 1939, declined to one seventh of the total by 1943.[7] In view of this fact, our total imports could increase only because our imports from other areas doubled and tripled. For example, our imports from Latin America increased from $518,000,000 in 1939 to $1,600,000,000 in 1944, while those from Canada increased from $340,000,000 to $1,300,000,000.[8] In view of the disruptive effects of war on international trade, it is difficult to foresee what the future international trade of the United States may be like.

QUESTIONS AND PROBLEMS

1. "The theory of interregional and international trade depends upon a considerable number of assumptions." Explain.

2. "The existence of interregional and international trade depends upon differences in the endowments of productive agents between regions and countries." Show whether you agree.

3. "Differences in the endowment of productive agents between regions or countries may be either absolute or relative." Explain.

4. "A region or country gains from trade only when it imports from other areas certain commodities which could not be produced at home." Show whether you agree.

5. "The Law of Comparative Advantage suggests that all regions or countries have an opportunity to gain from trade." Do you agree? Explain.

6. "When two regions or countries engage in trade, the prices of goods which enter into their trade tend to be the same in both areas." Show whether you agree.

7. "There is a tendency toward the equalization of the prices of the different varieties and grades of the productive agents between two trading regions or countries." Do you agree? Explain.

8. "The extent of the gain which a region or country realizes from trade depends upon the strength of its demand for the goods which it imports as compared with the strength of the demands of other areas for the goods which it exports." Explain.

9. "If the demand of one region or country for imported products increases, the terms on which it trades will tend to be altered in a manner which is unfavorable to that region or country." Explain.

10. Explain how the theory of trade is affected if we relax our assumptions concerning the mobility of the productive agents and the existence of equilibrium conditions.

11. How may trade between regions or countries be influenced by the existence of conditions of monopoly and monopolistic competition in many industries? Explain.

[7] *Survey of Current Business*, January, 1944, p. 13.
[8] *Survey of Current Business*, February, 1945, p. 21.

12. "If the agents of production are not freely divisible, trade may spring up between regions or countries where none would otherwise exist." Show whether you agree.

13. "The imports of any country or region must equal its exports." Do you agree? Explain.

14. Explain the method which we use in classifying the invisible items of trade as imports and exports.

15. "Trade statistics show that the exports of the United States almost always exceed the imports; hence the theory that exports and imports must be equal is clearly false." Show whether you agree.

16. If a country has a favorable balance of trade with respect to both commodities and services, how will this situation correct itself? Explain.

17. "Most countries and many smaller regions have attempted to limit and restrain their trade with other areas." Explain.

18. "The policy of restricting interregional and international trade is often based on the 'buy at home' fallacy." Discuss.

19. "There is bound to be more employment for labor within a given geographical unit if the people of that area produce all sorts of economic goods for themselves than if they rely on other areas for supplies of many goods." Show whether you agree.

20. Criticize the infant-industry and preparedness arguments for the restriction of trade.

21. "The principles of international trade are valid for all kinds of economic systems." Do you agree? Explain.

22. In 1940, the United States had an export surplus of $1,319,000,000 with respect to commodities and services. What were the offsetting items? Explain.

23. "The change in the character of our imports and exports and the change in our position from that of debtor nation to that of creditor nation have made it difficult for this country to maintain a favorable balance of trade in recent years." Show whether you agree.

24. "Many factors helped to account for the lowly status of the exports and imports of the United States between 1932 and 1939." Explain.

25. "Wartime conditions brought great increases in the extent to which the international trade of the United States was controlled by the federal government." Explain.

26. "World War II brought about a great decline in the international trade of the United States." Show whether you agree.

27. "The future status of the international trade of the United States is most uncertain." Why?

REFERENCES FOR FURTHER READING

Angell, J. W., *The Theory of International Prices*. Cambridge: Harvard University Press, 1926, Chaps. 14–18.

Griffin, C. E., *Principles of Foreign Trade*. New York: The Macmillan Company, 1924, Chaps. 1–7.

Huebner, G. G., and Kramer, R. L., *Foreign Trade Principles and Practices.* New York: D. Appleton-Century Company, 1942, Chaps. 1–8.

Ohlin, B., *Interregional and International Trade.* Cambridge: Harvard University Press, 1935, Chaps. 1–4.

Taussig, F. W., *International Trade.* New York: The Macmillan Company, 1927, Chaps. 1–13, 26–29.

Viner, Jacob, *Studies in the Theory of International Trade.* New York: Harper & Brothers, 1937, Chaps, 8, 9.

Young, J. P., *The International Economy.* New York: The Ronald Press Company, 1942, Chaps. 6–11, 15–18, 25–30.

XXVIII

Public Finance

In our discussion of the operation of the economic system we have often referred to various activities of government. It is now time to examine in some detail the activities which are carried on by government, the public expenditures which these activities necessitate, and the ways in which these expenditures are financed. "Government," it has been said, "as a form of social organization, has developed because, in the long run, it has afforded the means of supplying men with certain services more efficiently and more economically than these could have been supplied by each for himself." [1] All economic systems of the present day operate on the basis of specialization. Each individual performs a single function, or a limited number of functions, and depends on others to furnish him with the wide variety of commodities and services which he does not produce for himself. In the case of some services it seems the part of wisdom to look to the government for a more satisfactory performance than could be expected from private individuals or companies. The delegation to the government of the responsibility for providing or performing such services may be viewed simply as a part of the division of labor.

The functions which are performed collectively through government rather than by individuals are of several kinds. Some services, such as furnishing protection or maintaining law and order, would be very difficult, if not impossible, for the citizen to perform for himself. Others that could be and sometimes are carried out by private companies are often turned over to the government in the hope that they will thus be performed more efficiently or cheaply. Examples are the provision of water or electricity by municipalities. Still other functions could be performed by private individuals but probably would not be, since any individual's share in the resulting benefits would be rather small and long deferred. The establishment of institutions for dependents and defectives may be cited as an example. Even when an activity is left in the hands of private individuals, the government is often called upon to exercise a restrictive and regulatory influence. Thus

[1] H. L. Lutz, *Public Finance*. New York: D. Appleton-Century Company, Inc., 1936, p. 1.

we have a Pure Food and Drugs Act, a Sherman Anti-Trust Act, a Securities Act, and other legislation designed to protect the interests of the public.

PUBLIC EXPENDITURES

Expenditures of the Federal Government. The functions of governmental units, though they differ in many particulars, are all alike in one respect. They all involve the expenditure of funds. There has been a marked tendency for public expenditures to increase during the past few decades, not only in the United States but in other countries, and for both national and other governmental units. For example, the expenditures of the federal government of the United States increased from $517,000,000 in 1903 to $725,000,000 in 1913, $3,437,000,000 in 1928, and $12,774,000,000 in 1941.[2] The increase from 1903 to 1941 was 2,371 per cent. However, even the large expenditures of 1941 were smaller than those of some years at the time of World War I and very small indeed by comparison with those which were yet to come. During World War II, federal expenditures skyrocketed to a peak of $100,405,000,000, or about eight times the 1941 expenditures.[3]

Expenditures for protection (national defense) are always a very important item in the federal budget. In 1945, as we see in Table 24, direct expenditures for war amounted to $90,029,000,000 or 89.7 per cent of total federal expenditures. If expenditures for interest on the public debt and for veterans' pensions and benefits are included, the total cost of national defense amounted to 95.3 per cent of all federal expenditures.[4] In the prewar year 1941, expenditures for national defense amounted to 47.4 per cent of total expenditures, or 60.4 per cent if the expenditures for interest on the public debt and for veterans' pensions and benefits are included. Expenditures for such purposes as relief and work relief for the unemployed, social security, and aids to agriculture, which had amounted to 27.8 per cent of the total in 1941, were much less important in 1945. The cost of general governmental activities amounted to only 12.3 per cent of total federal expenditures in 1941 and 3.1 per cent in 1945.

Expenditures of State and Local Governments. The state governments of the United States spent a total of $207,000,000 in 1903, $359,000,000 in 1913, $1,753,000,000 in 1928, and $3,739,000,000 in 1944.[5] The increase from

[2] *Report of the Secretary of the Treasury, 1928,* p. 418, and United States Treasury Department statement of July 2, 1941.

[3] United States Treasury Department statement of August 1, 1945. All of these figures for federal expenditures are for fiscal years running from July 1 to June 30. Thus the 1945 fiscal year included the period from July 1, 1944, to June 30, 1945.

[4] The total cost of national defense certainly includes expenditures for the benefit of veterans, and would include all interest paid on the public debt of the federal government, if our national debt were large solely because the federal government, in certain past years, had spent for national defense at a more rapid rate than could be covered by the ordinary governmental revenues of those years. As a matter of fact, however, a part of the public debt resulted from governmental borrowing to finance heavy expenditures during the post-1929 depression.

[5] National Industrial Conference Board, *The Economic Almanac for 1945-1946,* New York: 1945, p. 106.

Table 24: **Expenditures of the Federal Government, Fiscal Years 1941 and 1945***

Type of Expenditure	Amount Spent in 1941 (*in millions*)	Per Cent of Total Expenditures, 1941	Amount Spent in 1945 (*in millions*)	Per Cent of Total Expenditures, 1945
National defense	$ 6,048	47.4	$ 90,029	89.7
Interest on the public debt	1,111	8.7	3,617	3.6
Veterans' pensions and benefits	553	4.3	2,044	2.0
Aids to agriculture	970	7.6	602	0.6
Social security†	943	7.4	982	1.0
Relief and work relief	1,637	12.8	5	—
General governmental activities‡	1,573	12.3	3,126	3.1
Debt redemption	64	0.5	—	—
Revolving funds (excess of credits, deduct)§	(125)	(1.0)	—	—
Total expenditures	$12,774	100.0	$100,405	100.0

* Source: United States Treasury Department statements, July 2, 1941, and August 1, 1945.

† In 1941, this item included expenditures of the Civilian Conservation Corps and National Youth Administration, as well as ordinary expenditures for social security.

‡ Includes expenditures for regular departments of government, reclamation projects, post office deficiency, refunds of taxes and duties, District of Columbia government, river and harbor work, flood control, and the Panama Canal.

§ Returns of funds from the Farm Credit Administration, Federal Farm Mortgage Corporation, and Public Works Administration.

1903 to 1944 was 1,706 per cent. The largest state expenditures in 1944 were for such items as education (28.0 per cent), highways (18.5 per cent), and public welfare (14.6 per cent). Less important items of expenditure included general government, public safety, natural resources, health, hospitals, correction, and interest.[6] Local governments in the United States spent $912,000,000 in 1903, $1,783,000,000 in 1913, $6,067,000,000 in 1928, and $4,554,000,000 in 1944.[7] Expenditures in 1944 were 399 per cent greater than those of 1903. Data covering the detailed items of expenditure by all local governments are not available, but the largest expenditures of cities with a population of over 25,000 are for education, public safety, public welfare, and highways.[8]

Causes of Increases in Public Expenditures. One rather obvious cause of increasing public expenditures in this country is the steady increase in the number of persons for whom governmental services must be performed. This growth in population, even if it called for nothing more than a duplication of existing governmental machinery, would bring about a propor-

[6] *Ibid.*, p. 131.
[7] *Ibid.*, p. 106.
[8] *Ibid.*, p. 131.

tional expansion of expenditures unless it were possible to achieve a decreasing cost per individual served. Moreover, the growth of population may cause an expansion of the needs which it appears necessary or wise to satisfy collectively rather than individually. The financial statistics of cities, for example, show a fairly steady per capita increase in public expenditures as we pass from the smallest cities to the largest.

Changes in the general price level and the purchasing power of money constitute another cause of the growth of public expenditures. The general price level in the United States was slightly higher in 1913 than in 1903, and was much higher in 1928 than in either of these other two years. The price level in 1939 was lower than in 1928, but in 1943 and 1945 it was considerably higher than in 1928. The national, state, and local governments, in order to render the services expected of them, must be able to obtain materials and supplies and in general to gain access to the agents of production much as does any businessman. The governments, in other words, must enter into competition with private individuals and firms for the use of the necessary agents of production. Hence, when the prices of goods and productive agents rise, governmental expenditures must increase also, even though there is no change in the number of functions performed or in the intensity with which they are carried on.

However, after due allowance has been made for growth of population, changes in the price level, and possibly other factors, the chief explanation of increasing public expenditures is to be found in the continually growing number of functions being performed by federal, state, and local governments. This expansion of functions in turn has been largely the result of such things as the development of great and pressing economic problems as our economic system increased in size and complexity, the growing conviction that social welfare can and should be advanced by collective action through the government, and the more modern notion that the government can and should assume direct responsibility for the satisfactory operation of the economic system as a whole and for the maintenance of a high level of income, production, and employment.[9]

Economic Problems and Public Expenditures. With the development of our economic system, production has become highly specialized, roundabout, and large scale. The economy has become increasingly dependent upon a smoothly functioning system of exchange, and economic interdependence has increased. As the economic system became larger and more

[9] One other factor aiding and abetting the expansion of governmental functions and expenditures has undoubtedly been the popular attitude toward public funds. Many people seem to regard the public treasury as a vast reservoir into which flow mysterious but inexhaustible streams of wealth. It appears to such people that the problem in connection with the treasury is not one of economy in appropriations and expenditures, but rather that of seeing to it that their particular district or locality receives its full share of the funds—or more than its share if possible. This attitude is, of course, invalid. The government gets its money income from the money incomes of the citizens.

complex, problems developed which often caused distress to large numbers of people and seemed to call for collective action. Examples are problems of money and banking, public utilities, transportation, monopolies and trusts, labor relations, and the issuance of securities. In many cases the attempted solutions have involved governmental regulation, which has not required large public expenditures though it has added greatly to the list of governmental functions. The citizens of the economy continued to hold in general to the capitalistic principle that governmental interference with and control over the economic activities of private citizens is in itself not a good but an evil. However, they were willing to embrace this evil in cases in which it appeared to be the means of preventing or eliminating a greater evil which would otherwise exist. Governmental regulation in connection with such individual economic problems did not keep the economy as a whole from operating on a capitalistic basis.

Welfare Activities and Public Expenditures. State and local governments have come to spend large sums for highways, education, and public welfare in general. Such expenditures are thought to make for the continued development of our people and their economic activities, and hence for the general welfare. When the automobile was first introduced, highway construction and maintenance could be provided at a cost which would now seem remarkably low. However, the use of the automobile for pleasure has increased so rapidly that a car is at present regarded by most families as a necessity, while its use for commercial purposes has grown about as rapidly. Traffic on the main highways has become extremely heavy, and the expense of constructing and maintaining macadam and concrete surfaces which will support giant motor trucks is very high.

As an economic system increases in complexity, it becomes more and more apparent that adequate preparation for one's life work is essential to the achievement of a high measure of success, and it becomes extremely difficult for a person to rise from the ranks if he depends upon his native ability unaided by occupational training. Also, as standards of living improve, greater educational facilities are demanded and obtained. For these and other reasons we are at present attempting by collective action to provide everyone with a certain minimum amount of training, and to assist as many as possible to pursue higher education. Since little or no effort is made to place most of our educational activities on a self-supporting basis, the cost in the form of public expenditures is large.

There have been other large increases in expenditures for hospitals, clinics, health and sanitation, recreation, and institutions for delinquents, dependents, and defectives. In the dreary days when the insane were left to "mumble in the chimney corner," when the feeble-minded were allowed to roam at will and were merely regarded as a bit queer, when sickness or even plagues were thought to be marks of divine disfavor or the work of

evil spirits rather than the result of lack of sanitation, and each family was expected to pay for its own medical attention or to go without, governmental expenditures for such purposes were not large. But nowadays we regard the satisfactory handling of these matters as vital to the public welfare, and undertake to provide for them through governmental agencies. Provision of this kind is praiseworthy, but it is also expensive. However, the welfare activities of these kinds do not tend to interfere significantly with the operation of our economy on a capitalistic basis.

Governmental Functions and Expenditures in the Great Depression. The idea that the federal government should assume direct responsibility for the satisfactory operation of the whole economy came to the fore in the depression years following 1929. The United States had many business depressions prior to 1929, but the earlier depressions had been allowed to run their course, apparently on the principle that recovery attained in the "natural" manner would be more prompt and lasting. But in the post-1929 depression there was a widespread demand that the government take action both to promote recovery from the depression and to relieve the millions of unemployed and destitute. The response of government to this demand resulted in greatly expanding functions and large public expenditures.

The activities of the federal government in providing relief or work relief for the unemployed, through the Public Works Administration, the Civil Works Administration, the Civilian Conservation Corps, the Work Projects Administration, and assistance given the states in providing direct relief, were very costly. From 1933 to 1941, inclusive, the total expenditures for unemployment relief were about 18 billion dollars, and amounted to about 26 billion dollars if public works are to be included as a part of unemployment relief. Again, the federal government, through the Agricultural Adjustment Administration, the Farm Credit Administration, the Commodity Credit Corporation, and other agencies, attempted to regulate agricultural production, raise the prices of farm products, refinance farm mortgages, and lend the farmers credit on their holdings of various crops. The expenditures for these purposes were more than a billion dollars in some years, and totaled over 5 billion dollars for the period from 1933 to 1941, inclusive.

Other federal activities designed to promote recovery were less costly than those already mentioned. They included financial assistance to railroads, banks, and other institutions through the Reconstruction Finance Corporation, the refinancing of the obligations of homeowners other than farmers; the promotion of the rehabilitation of industry and business through self-regulation under the National Recovery Administration; insurance for bank depositors through the Federal Deposit Insurance Corporation; the Subsistence Homestead Projects; and the Emergency Housing Program. At the same time the federal government undertook other activi-

ties which seemed to be dedicated to reform rather than, or in addition to, recovery from the depression. Examples included the regulation of the issuance of new securities and of the activities of security exchanges through the Securities and Exchange Commission under laws of 1933 and 1934, the regulation of the public utility industry through the Securities and Exchange Commission and the Federal Power Commission under the Public Utilities Act of 1935, and various activities for the benefit of labor carried on under the National Labor Relations Act of 1935 and other laws.[10]

The War Period. Most of the depression activities of the federal government were still going full blast when World War II began. When the United States entered the war, governmental control over the economic activities of the country increased greatly and soon dwarfed anything that had ever existed before in this country. Long before the war was over, agencies of the federal government were controlling output in many branches of production, the prices of commodities and services, wages and salaries, rents, the allocation of essential materials and equipment among industries and businesses, the allocation of labor among industries, businesses, and the armed forces, industrial relations in many industries, imports and exports, the apportionment of certain scarce consumers' goods among the individual citizens, and many other things. While many people wondered whether all the governmental controls were essential to the successful prosecution of the war and whether some of them were well suited to the objectives which were being pursued, in general the wartime functions and expenditures of the federal government were accepted with good grace by the people as being more or less inevitable in such an emergency period.[11]

Proposed Governmental Functions. During the war period many people in this country apparently came to the conclusion that, while some controls over economic activity imposed by the federal government in wartime should be relaxed in the postwar period, the federal government must continue to assume responsibility for the successful operation of the economy as a whole. Governmental functions proposed in this connection include the guarantee or underwriting of full employment by the government, the provision of a system of social security, popularly known as the "cradle-to-the grave" variety, which will cover many more people, provide against more risks, and furnish much larger benefits than the present system; the provision of high minimum wages or even a national minimum per capita income; and support for a high wage policy in general in order to maintain

[10] In practice, the distinction between recovery and reform activities was not at all clear cut. Many activities dedicated to recovery also included certain reform features, and it was doubtless hoped that some of the reform activities would also be of some assistance in promoting recovery from the depression.

[11] The wartime activities of the federal government will be examined in greater detail in Chapter XXXI.

purchasing power and the total demand for goods. Such functions, as we shall see later in the present chapter and in the chapters on economic instability which follow, would have most serious implications for the future of our economic system.

Proper Activities of Government. Since the size of public expenditures depends to a great extent upon the number and cost of the services which various governmental units are called upon to perform, it is important to inquire how many functions should be assigned to government. Some functions must be performed by the government if they are to be performed at all, but others could be carried on either by government or by private individuals and firms. In the latter cases, a function should be delegated to the government only when it appears that, bearing in mind the necessary costs of administration in collecting and disbursing the funds together with whatever knowledge may be available as to the efficiency of the government in question, the expenditure of a given sum collectively will result in a more adequate and economical service than would be obtained by a similar sum privately spent. However, this tells us nothing concerning the aggregate number of functions which should be performed by government.

A general decision concerning the aggregate of governmental functions and expenditures may be reached by having recourse to the familiar economic terms "satisfaction" and "productivity." Individuals spend their money incomes for consumers' goods, usually after a consideration of the different amounts of satisfaction that are likely to be derived from the various alternative uses to which the funds in question might be put, or for producers' goods after a similar decision has been reached with regard to productivity. In other words, individuals try to spend their incomes in such ways as will be likely to result in the realization of the greatest possible amount of satisfaction, or productivity, as the case may be. Applied to public expenditures, the principle as commonly stated is that additional funds should be collected and spent by the government only so long as the amount of satisfaction to be derived in the aggregate from the new governmental functions or services is likely to exceed the amount lost in the aggregate by having to give over the funds in question to the government.

This principle should be modified to allow for the fact that funds may be spent and hence have to be collected in such amounts and in such ways as will indirectly have the rather serious effect of restraining or depressing business. Hence we may say that a government should collect and spend such an aggregate sum that the advantage in the form of satisfaction to be gained by any further expenditures and functions will not be sufficient to offset the disadvantage of loss of satisfaction, either direct, or indirect through the ill effects on business, by reason of turning the funds in question over to the government. Furthermore, just as an individual attempts

to distribute his income among the various objects of expenditure in such a way as to make the last dollar spent for each purpose afford him as much satisfaction as could be derived from a like expenditure for any other object, so too it might be desirable to have the aggregate expenditures of government split up among the several governmental activities in such a manner that no money will be spent for one purpose which would result in greater satisfaction if it were added to the sum to be spent for some other purpose. There is little chance that these principles could be applied with great precision in practice, but they are the ones which should be borne in mind.

Evaluation of Governmental Functions and Expenditures. On the basis of the main principle proposed above, does it appear that certain activities of our several governmental units should be curtailed or dropped entirely? We may first express doubt that any net social advantage would result from a decision not to have the governments carry on the various functions which they were performing in the regulation and promotion of economic activity prior to 1929. Nor does it appear that greater satisfaction would be created by leaving in private hands the funds which are collected and spent by governmental units for certain social welfare purposes. The need for these services has come upon us gradually and inevitably with the changes that have taken place in our social structure, and it does not seem that the responsibility for performing these functions could have been avoided under any circumstances. Under present conditions it is unthinkable that such matters as education, highways, and the provision of institutions for dependents, defectives, and delinquents should be neglected. On the whole, then, it must be said that the great bulk of these activities of governments may be expected to stand up under the scrutiny of the proposed test for public functions and expenditures. That is, these activities probably create satisfactions for the receivers of the services, or for society as a whole, which more than compensates for the loss of satisfactions to individuals or damage to business which may result because funds are collected in order to make it possible to carry on these functions.

There is ordinarily more doubt whether a similar conclusion can be reached with regard to what has been the principal activity of the federal government—the provision of protection or national defense. Expenditures for this purpose, as has been mentioned, make up a very large part of the total expenditures of the federal government over the years, and consist in the main of payment for past wars and preparation for those of the future. Without attempting to ascertain whether the wars of the past were necessary or the extent to which maladministration may have added to the costs of these conflicts, we may note that there is a conviction on the part of many persons that, under ordinary circumstances and from the economic point of view, war is an unwise and costly method of "settling" disputes

between nations. Modern wars almost inevitably bring economic losses, and not economic gains, to all the participating nations.

If the principle of satisfaction is applied to this type of expenditure, there seems to be only one way in which the principle may be stretched sufficiently to make it cover the activity. If the conditions of the world in the recent past are regarded as fixed, unyielding conditions to which we may react but which we cannot change, then preparation for war, and at times the actual waging of war, may be necessary. We do not want aggressor nations to be in a position to take our land and resources from us, or to destroy our institutions. To prevent such heavy losses, it may be worth while for us to bear the enormous burdens involved in preparing for and waging war. Thus war may be the lesser of two great evils, even if it results in severe economic losses. On the other hand, from the point of view of the world as a whole, the hundreds of billions of dollars spent for wars and armaments in the past have been economic waste. There would be a vastly greater sum total of satisfactions from the use of economic goods if the world were freed from military conflict.

What shall we say of the depression activities of the federal government? From the point of view of a direct comparison of satisfactions, many of them would probably pass muster. Consider, for example, relief or work relief for the unemployed. In so far as governmental assistance was received by those who really needed it, it is hard to believe that these expenditures did not give more satisfaction to the recipients of the government's aid than they took away, or will take away, from the taxpayers. However great the pangs suffered by those who have paid or must pay taxes to cover relief expenditures, they are doubtless less painful than the pangs of death by starvation or exposure.

However, when we consider the indirect effects of the depression activities and expenditures, the answer is not so clear. The heavy expenditures had to be financed either by taxation or by borrowing. To the extent that taxation was used, high taxes tended to cut into employment-creating private spending for consumption or investment as well as into unnecessary saving. Moreover, some individuals in the higher tax brackets may have been rendered unwilling to put capital funds into risky ventures, since any gains which they might make would have to be shared with the government, while any losses which they might suffer would be entirely their own. To the extent that borrowing (deficit spending) was used to finance the governmental expenditures, business confidence may have been affected and fears of higher taxes later on aroused, with adverse effects on the amount of production and employment furnished by private industries and businesses.

Funds for relief or work relief may have been used in some cases to secure political objectives rather than to promote recovery from the depres-

sion, and the receipt of governmental assistance undoubtedly weakened the morale of some persons and made them anxious to make a "career" of the WPA or to "retire" on relief, rather than to shift for themselves in private employment. The laws of 1933 and 1934 may have cleansed and purified the issuance of securities and the operation of the security exchanges, but, according to some authorities, they also brought the issuance of new corporate securities almost to a standstill and greatly weakened the functioning of the security exchanges. The activities in behalf of labor may have greatly improved the bargaining power of employed workers while making enterprisers less willing than they would otherwise have been to undertake production and furnish employment. Direct competition with private business by governmental enterprises or agencies, or even the threat of such competition, and the general attitude of hostility toward private industry and business on the part of the federal government, may have had similar effects.

Other examples could be given, but those which have been presented may be adequate to suggest that the assumption by the government that the capitalistic system would not operate successfully may have led to governmental activities which tended to make it relatively sure that the capitalistic system would not operate successfully. It is difficult to decide whether the economic system as a whole was actually better off by, say, 1940 than it would have been if the federal government had not assumed responsibility for the satisfactory operation of the system and for the promotion of recovery in 1933. However, it is certain that the degree of recovery from the depression which occurred under governmental control was rather disappointing and that the recovery was far from complete at the outbreak of World War II.

In evaluating governmental activities which are proposed for the postwar period, such as the guarantee of full employment, the provision of cradle-to-the-grave social security, and the underwriting of high minimum wages or minimum per capita incomes, we could not be content simply to compare the satisfactions gained by the beneficiaries of these policies with losses of satisfaction on the part of taxpayers. As in the case of the depression activities and expenditures, we would need to consider the effects of these activities on private businesses and industries. If, as some people think, the enthusiastic pursuit of the policies in question would make it impossible for our economic system to operate in capitalistic fashion and would result in its transformation to some form of controlled and planned economy, we would have a great deal to worry about. It would be necessary to consider the loss of economic freedom which we would experience in a planned and controlled economy, the possible loss of our political freedom as well and the extent to which we would miss it, the difficulties which might arise because of the lack of adequate incentives in the planned

economy, and other matters. Clearly it would not be a simple problem to evaluate the proposed activities of the federal government.[12]

TAXATION

Once it has been decided what functions shall be performed by governmental agencies, there remains the task of determining the method or methods of procuring the funds from which the proposed public expenditures may be made. At times national governments have deemed it necessary to manufacture their own purchasing power. In other words, they have printed large issues of inconvertible or "fiat" money which they have used in making their expenditures. The effect of this additional purchasing power, competing with the money already in circulation for the limited existing supplies of commodities and services, has been to deprive the individual citizens of a part of their purchasing power just as effectively as if the governments had exacted tribute from the citizens in the first place.

Governments also commonly borrow from the citizens or the banks to obtain funds for public expenditures. This method has been used when it has appeared undesirable or impossible to obtain sufficient funds for current expenditures from current sources of revenue. By borrowing, we get governmental services in the present and pay for them on the installment plan over a long period of years. That is, if interest and principal are to be paid on government bonds, borrowing must be regarded merely as deferred taxation. Occasionally governments secure small amounts of direct revenue from the operation of various public enterprises and from other nontax sources, but these items are not usually very important.

The Nature of Taxation. Taxation must be regarded therefore as the long-run source of almost all the funds from which public expenditures are made. A tax is "a compulsory contribution from the person to the government to defray the expenses incurred in the common interest of all, without reference to special benefits conferred."[13] In connection with this definition several observations should be made. First, a tax is a compulsory contribution in that the amount to be paid is decided by the government, as are also the time and method of payment, without regard for the wishes of the individual taxpayer. Second, there is in taxation no definite *quid pro quo;* that is, the taxpayers are required to contribute to the support of the government on some basis other than the amount of benefit or service received by them from the government. Finally, the purpose of taxation is to provide revenue for carrying on various functions in the interest of the public. In this connection a serious question often arises as to whether taxation may properly be made the instrument for the accomplishment of some

[12] For a more extended discussion of these governmental activities which are proposed for the postwar period, see Chapter XXX.

[13] E. R. A. Seligman, *Essays in Taxation.* New York: The Macmillan Company, 9th ed., 1921, p. 432.

ulterior purpose, such as the reduction of existing inequalities of wealth and income.

General Principles of Taxation. In view of the main purpose of taxation, the first and most important test of a tax system is whether it will furnish sufficient revenue. The satisfaction of this test alone does not insure a perfect system of taxation, for many other important considerations must be borne in mind; but the inability to meet this test is of itself quite enough to make the tax system a failure. A second principle of taxation is economy in collection. Other things being equal, taxes which involve the collection of large increments of revenue, with a minimum of complaints and bookkeeping details to be handled, are preferable to those which necessitate the collection of a multitude of small sums and involve great complexity of administrative machinery. In the words of an early economist, "every tax ought to be so contrived as both to take out and keep out of the pockets of the people as little as possible over and above what it brings into the public treasury of the state."

Simplicity is another desirable feature of a tax system. The provisions of our tax laws should not be complicated, but should be so worded that they may be readily understood by those who must pay the taxes as well as by those who administer the tax laws. A sound tax system will also be characterized by certainty. Our economic system can adjust itself to a fairly high or a low level of taxation with something like equal success in the long run, but uncertainty as to what taxes will be seems ever to be productive of undesirable results. Taxes should be collected at such times and in such a manner as will be most convenient for the taxpayers, and the tax system should be elastic so that it can be made to yield increased or decreased revenue as the needs of government change.

A final requisite for a sound tax system is that it shall be equitable. Once it is ascertained that a given system of taxation is likely to provide an adequate amount of revenue, the next most important consideration is to have the tax burden distributed among the taxpayers in an equitable manner. This consideration becomes increasingly significant as the size of the aggregate burden of taxation increases. Years ago, when public expenditures were relatively slight and the total burden of taxation was not heavy, almost any system was likely to be reasonably fair. This would not be true at the present time. In view of the importance of the principle of equity, it will be well to inquire further into its meaning as applied to a system of taxation.

The Principle of Ability to Pay. Most economists agree at the present time that equity in taxation requires that the tax burden should be distributed on the basis of ability to pay. That is, each individual should be taxed according to his ability to make contributions to the support of the government, without regard for the amount of benefit which he may derive

from the activities of the government. This principle seems on the surface to be admirably simple, but its interpretation is extremely difficult. What, for example, shall be the test of ability to pay? The answer to this question used to be that the possession of wealth was a clear indication of ability to pay taxes, and it was decided that any man who was fortunate enough to own great quantities of wealth should be required to contribute large sums to the governmental treasuries. As time went on, economists came to suspect that some modification of this answer was necessary, for many individuals were receiving large incomes annually but were escaping the tax-gatherer almost completely under the prevailing system of general property taxes.

Today the greatest emphasis is placed upon the receipt of income as the indicator of ability to pay taxes, although the ownership of wealth is by no means completely disregarded in this connection. This does not mean, however, that it is necessary merely to discover the amount of a person's income in order to pass judgment on his ability to pay taxes. Many other matters should be borne in mind, such as whether the income in question is "earned" or is derived from the ownership of wealth, whether it contains any elements such as the return on diminishing assets, whether it represents any surplus over and above the returns necessary to induce the individual to continue to perform his services or lend his capital, and, finally, the length of time over which the income is received.

Progressive, Proportional, and Regressive Taxes. Probably the most important question in connection with the principle of ability to pay relates to the way this ability increases as income increases. In other words, as a man's net taxable income increases, does his ability to pay taxes increase exactly in proportion to the increase in income, or more or less rapidly? If A has a net taxable income of $10,000, while that of B is $5,000, is A able to pay exactly twice as much in taxes as B, more than twice as much, or less than twice as much, supposing the considerations mentioned in the preceding paragraph to be the same for both individuals? If we decide that A's ability to pay is exactly twice as great as B's, then taxation should be proportional—that is, the same rate should apply to both individuals—and at this rate A's contribution will be double that of B. If A has less than twice the ability of B to pay taxes, the rates of taxation should be regressive; that is, a lower rate should apply to the larger income than to the smaller. Finally, if A's ability to pay taxes is more than double that of B, the rates of taxation should be progressive; that is, the larger the income, the higher the rate of taxation which should be applied to it.

The advocates of progressive taxation base their argument in large measure on the theory that the satisfaction to be derived from the expenditure of a unit of money income diminishes as the total income of the spender increases. A certain amount of income is necessary as a minimum

for subsistence, and does not represent ability to pay in the true sense of the term. The first increments of income above this minimum amount are used by the individual to satisfy wants which are fairly important, but further and yet further additions to income will be used to satisfy less and still less important wants. It is said to follow that, unless we are dealing with individuals who choose to satisfy their less pressing desires before attending to those which are more urgent, diminishing satisfaction is experienced in the expenditure of successive increments of income. Thus it is held that the man with an income of a million dollars a year is not losing nearly so much satisfaction when compelled to give up one hundred thousand dollars in taxes as is the man with ten thousand a year when he contributes a thousand dollars to the support of the government, although it is clear that the rate is the same in these two instances. It may seem equitable, then, for the recipients of large net taxable incomes to pay taxes at higher rates than those applied to the smaller incomes of other men, which means, of course, the use of progressive rates of taxation.

Since there is no way to make accurate direct comparisons of utilities or satisfactions between different individuals, and since it is by no means certain that all individuals have equal capacities for experiencing satisfactions, it is impossible to *prove* that this analysis in support of progressive taxation is sound. However, most students of public finance are satisfied that progressive rates of taxation should be used wherever they can be successfully applied—which means for all practical purposes in the taxation of incomes and inheritances. It should be apparent that a progressive tax upon a commodity, say cigarettes, which would mean a high tax if the purchaser were rich and a low tax if he were poor, would not be practicable. Once it is decided that taxes should be progressive where possible, there remains the problem of deciding how progressive they should be, and here we must consider the indirect as well as the direct effects of taxation. That is, taxes might be made so highly progressive that they would discourage business activity, curb individual initiative, and check unduly the accumulation of capital.

The Incidence of Taxation. In deciding whether a particular tax system is equitable, it is of vital importance to know who ultimately bears the burden of the various taxes levied. It has long been customary to classify taxes into two groups, direct and indirect. Direct taxes are those which are collected at the outset from the persons on whom it is intended that the burden shall fall. Indirect taxes are those collected from one group of individuals with the expectation that the burden will be shifted by them to a different group. This classification of taxes into direct and indirect groups has proved to be unfortunate. In some cases, we find so-called direct taxes being passed on, in part at least, to others than those from whom the taxes were originally collected, while at other times so-called indirect taxes have

not been shifted, but have remained a burden upon those who paid them in the first place. Wholly apart from any classification, however, the problem of the "incidence," or ultimate burden, of taxation is extremely important.

The study of the shifting and incidence of taxation is a branch of the study of value. To know whether a tax on a given commodity is likely to be shifted, it is necessary to ascertain whether the conditions of production and sale for that commodity make it possible for its price to be raised so as to pass the amount of the tax on to the consumers or for the tax burden to be passed backward to the suppliers of certain productive agents in the form of lower prices. If the tax in question is levied on some agent of production, such as land or capital, it is necessary to make a similar investigation into the conditions which determine the price of this agent of production, in order to discover whether the tax can or cannot be shifted.

The Federal Revenue System. As may be seen in Table 25, the federal government had total tax revenues of $44,170,000,000 in the fiscal year 1945, and total revenues from all sources of $47,639,000,000. The comparable figures for 1941 were $7,761,000,000 and $8,581,000,000. Of the total revenues in 1945, the personal income tax yielded 39.5 per cent; all corporation taxes 34.8 per cent; payroll taxes, 3.8 per cent; estate and gift taxes, 1.3 per cent; excise taxes, 12.5 per cent; and customs duties, 0.8 per cent. These figures may be compared with those for 1941, when the personal income tax yielded 16.5 per cent; corporation taxes, 25.8 per cent; payroll taxes, 10.8 per cent; estate and gift taxes, 4.7 per cent; excise taxes, 28.0 per cent; and customs duties, 4.6 per cent. We shall now turn to an examination of these several types of taxes.

The Personal Income Tax. The taxation of personal incomes was the most important source of federal revenue in 1945. In the application of the tax, after the deduction from gross income of the necessary expenses of acquiring it and certain other allowable deductions, an exemption of $500 was allowed the individual taxpayer against the normal tax, which was applied at the rate of 3 per cent to the rest of his income. Against the surtaxes, or additional taxes which made the personal income tax progressive, the individual taxpayer was allowed an exemption of $500 for himself and $500 more for each dependent. The surtaxes began at 20 per cent on the first $2,000 of eligible income, and reached a maximum of 91 per cent on all income in excess of $200,000 above the exemptions. The total of normal tax and surtax was limited to 90 per cent of net income.[14]

In 1945 normal taxes and surtaxes on wages and salaries of $5,000 or less were withheld by the employers and paid directly to the government.

[14] The Revenue Act of 1945, effective January 1, 1946, changed the personal income tax to some extent. Exemptions for taxpayers and their dependents are now as large for normal tax purposes as for surtax purposes, surtax rates are 3 points lower than formerly in each bracket, and each taxpayer is required to pay only 95 per cent of his total tax, as computed.

Table 25: Sources of Federal Revenue, Fiscal Years 1941 and 1945*

Source of Revenue	Amount Received in 1941 (*in millions*)	Per Cent of Total Revenue, 1941	Amount Received in 1945 (*in millions*)	Per Cent of Total Revenue, 1945
Personal income tax.....	$1,418	16.5	$18,831	39.5
Corporation income tax..	1,852	21.6	5,083	10.6
Excess profits taxes.....	201	2.3	11,148	23.4
Capital stock tax	167	1.9	372	0.8
Payroll taxes.....	926	10.8	1,793	3.8
Estate and gift taxes.....	407	4.7	643	1.3
Excise taxes.....	2,399	28.0	5,945	12.5
Customs duties	391	4.6	355	0.8
Nontax revenue.....	820	9.6	3,469	7.3
Total revenue	$8,581	100.0	$47,639	100.0
Transfers to trust funds (deduct)....	854		1,283	
Total net revenue...	$7,727		$46,356	

* Sources: *Federal Reserve Bulletin*, October 1945, p. 1047, and United States Treasury Department statements of August 18, 1941, and August 1, 1945.

Individuals with incomes of more than $5,000 from wages and salaries, or with incomes of more than $100 from other sources (provided total income was at least $500), had to file an estimate, by March 15, 1945, of their total tax liability, the amount of tax which would be withheld from them, and their net tax liability to the government. One fourth of the latter sum had to be paid at once, and the remainder in three installments on June 15 and September 15, 1945, and January 15, 1946. Then on March 15, 1946, all payers of the personal income tax had to file a final return adjusting their actual tax liability to the payments which had been made during the preceding year by withholding or otherwise.

Merits of the Personal Income Tax. The personal income tax is generally considered to be a good tax. It falls directly on income, from which all taxes must come eventually, if sometimes indirectly; and income is the most widely accepted indicator of ability to pay. The tax usually yields a large revenue but, being based on realized income, the receipts are likely to decline sharply in times of business depression when the government needs especially large revenues. The tax lends itself readily to progressive rates, which are usually thought necessary to the application of the principle of ability to pay. It is certain as to time and manner of payment, but it is only moderately satisfactory from the point of view of economy in collection.

Defects of the Personal Income Tax. Although repeated attempts have been made to simplify income tax procedure for the small taxpayers, the personal income tax law remains complicated, and a person of considerable

income often needs legal advice in arriving at the amount of tax he must pay. Except for the withholding of the tax on wages or salaries of $5,000 or less, the method of assessment is by the taxpayer's declaration of his income, expenses, and deductions, supplemented by information as to amounts paid in salary, interest, or other types of income, furnished by employers and others who make the payments. This method of assessment requires a high degree of administrative efficiency if the tax is not to be merely one on honesty and to lead to widespread evasion.

It is difficult to define income for purposes of taxation, and our laws do little more in this respect than to enumerate various taxable and nontaxable items. Under the law, as now interpreted, some peculiar situations arise. For example, a farmer need not count as taxable income the food and shelter provided by his farm, but no similar privilege is accorded those who must buy their food and shelter with money income. A houseowner who occupies his house need not count as taxable income the shelter he enjoys, but if he rents the house to another, the rental that he receives is taxable. If a man is buying a house through payments over a long period of time, the interest that he pays on his indebtedness is deductible from income for tax purposes, but he has no such deduction if he buys the house outright for cash. If the individual owns a corporate bond, the interest received is subject to the income tax, but the same thing is not true if he owns a "tax-exempt" bond issued by some governmental unit in the past.

The most serious defect of the personal income tax, under present highly progressive rates, is its effect in discouraging enterprise and personal initiative. In combination with taxes on corporate income, it imposes heavy double taxation on income derived from corporate ownership. By itself, it discourages self-employment by individuals and the foundation of unincorporated enterprises. The income from business ownership is always uncertain, and this is especially true of the income from new ventures, yet the personal income tax takes a considerable part of any net income received while permitting only a relatively meager offset for losses. The tax is especially restrictive for individuals who already have large incomes. Why should the motion-picture star undertake another picture this year if all the additional income derived from it will be subject to a 91 per cent tax, and why should the well-to-do individual undertake a new and risky business venture under the same circumstances?

The Incidence of the Personal Income Tax. The usual conclusion with regard to the incidence of the personal income tax is that its burden remains on those who pay it and cannot be shifted. It is argued that there is nothing about the income tax that enables individuals or companies to raise the prices of the goods they are selling, or that makes the personal services or the capital furnished by individuals command a higher return. Consequently, it is said, all efforts to pass personal income taxes on to others

are doomed to failure. This is undoubtedly true in so far as the individual taxpayer is concerned. However, if enough individuals were discouraged from undertaking business ventures or entering high-paying occupations, those who still did these things might get large enough incomes so that, after paying the tax, they would have as much income left over as they had in a previous situation in which the tax had not yet been imposed.

Taxes on Corporations. The corporation income tax produced 10.6 per cent of total federal revenues in 1945. The normal tax rate was 24 per cent for corporations with net incomes in excess of $25,000, and 15 to 19 per cent for corporations with smaller net incomes.[15] The surtax was 16 per cent for corporations with net incomes in excess of $50,000. Corporations with net incomes between $25,000 and $50,000 paid 10 per cent surtax on the first $25,000 and 22 per cent on the balance, while corporations with a net income of less than $25,000 paid a straight 10 per cent surtax.[16]

Excess profits taxes produced 23.4 per cent of total federal revenues in 1945. Most of this revenue was yielded by the excess profits tax applied in 1940 (sometimes called the "main" excess profits tax to distinguish it from the "declared value" excess profits tax). For the purposes of this tax, each corporation received a flat exemption of $10,000 plus a further credit computed by either of two methods—net income or invested capital. By the net income method, the credit amounted to 95 per cent of the average net income of the corporation in the period 1936–1939, plus 8 per cent of net capital additions or minus 6 per cent of net capital reductions. By the invested capital method, the credit was 8 per cent of invested capital up to $5,000,000 and 7 per cent of invested capital over $5,000,000. After the deduction of the exemption and the credit, each corporation (unless specifically exempted by law) paid the excess profits tax on the remainder of its income at the rate of 95 per cent. Ten per cent of the tax was to be refunded after the war.[17]

In addition to the taxes already described, the capital stock tax and declared value excess profits tax fell on corporate incomes in 1945, though they produced relatively little revenue. These taxes had been paired as complementary taxes since 1933. The rate of the capital stock tax was $1.25 per $1,000 of the declared value of each corporation's capital stock. Having declared a total value for its capital stock, a corporation could earn 10 per cent on this value without paying any declared value excess profits tax. However, it had to pay 6.6 per cent on net earnings of 10 to 15 per cent, and 13.2 per cent on net income in excess of 15 per cent, on the total de-

[15] The normal tax applied only to corporate net income which was not subject to the main excess profits tax.

[16] The Revenue Act of 1945 reduced surtaxes on corporate income by amounts ranging from 2 per cent for corporations with net income in excess of $50,000 to 4 per cent for corporations with net income of less than $25,000.

[17] The main excess profits tax was repealed by the Revenue Act of 1945, effective January 1, 1946.

cleared value of its capital stock.[18] Under these various taxes on corporations, as described, it would seem that a corporation might have to pay more than 100 per cent of its net income in taxes to the federal government, but such disasters were forestalled by a general proviso that all corporate taxes together must not take over 80 per cent of any corporation's net income.

Evaluation of Taxes on Corporate Income. The corporation income tax has usually been a good revenue producer although, like the personal income tax, its yield diminishes greatly in poor business years. This tax is not open to some of the objections urged against the personal income tax, though in the case of both it is difficult to decide what constitutes net income. Excess profits taxes also amount to very little in years when the level of economic activity is low or moderate, but they produce large amounts of revenue in prosperous years, and especially in wartime, when they are useful in preventing war profiteering. Excess profits taxes are often complicated, for it is very difficult to make them both fair and simple. They often give rise to extensive litigation and impose a severe test on the federal administrative machinery.

The various taxes on corporate incomes, considered together, are subject to certain criticisms, noteworthy among which is the suggestion that they do not conform to the principle of ability to pay. Many people contend that a corporation has no ability to pay taxes apart from the ability of its stockholders to pay. If this is true, we must note that the size of a corporation's net income has no necessary relationship to its stockholders' ability to pay taxes. One corporation with a moderate net income may distribute it in large amounts to its few (and possibly well-to-do) stockholders, while another with a huge net income may pay this income in small driblets to its hundreds of thousands of small stockholders. In such a case, a small tax on the former corporation's income and a heavy tax on the latter's will not cause the respective stockholders to give up income in accordance with the principle of ability to pay.

It is also commonly held that taxes on corporate income, in combination with the personal income tax, lead to double taxation of the same income. When the corporation receives net income, it pays the corporation income taxes, and any other levies on corporate income which may be in effect at the time, and then pays dividends to the stockholders. The dividends then become personal income to the stockholders and are subject to both the normal personal income tax and the surtax, if the stockholders' incomes are sufficiently large. This is unquestionably double taxation and it penalizes severely the income drawn from corporate ownership as compared with other types of personal income. This criticism still applies, though with

[18] The capital stock tax and declared value excess profits tax were repealed by the

somewhat reduced force, now that the excess profits taxes have been abolished.

The Incidence of Taxes on Corporate Income. The above argument concerning double taxation assumes that taxes on corporate income cannot be shifted forward to consumers of corporate products or backward to the owners of productive agents used by corporations, but remain as a burden on the corporation and its owners. This assumption is probably valid for most practical purposes since the taxes in question fall on the net income realized from business operations in a given period and do not affect directly the marginal costs of the firms or give them any direct ability to raise prices. If any shifting of the tax occurred, it would be over long periods of time and in industries made up largely of corporations and not severely regulated by the government. The process would lower the attractiveness of investments in corporate enterprises and reduce investment, with the result that surviving corporate firms could get higher prices and possibly as much net income as before the heavy taxes were applied. Any shifting that occurred in the long run would not, of course, alter the fact that double taxation had existed in the meantime.

Payroll Taxes. In 1945, payroll taxes produced 3.8 per cent of the revenues of the federal government. These taxes are collected for the accumulation of reserves required under the Social Security and Railroad Retirement Acts. Under the former act, both employers and employees pay 1 per cent of wages up to $3,000 a year in connection with the old-age annuity program, while employers pay 3 per cent on their payrolls for the support of the unemployment program. The taxes paid by employees operate as a crude sort of income tax, since all employees who come under the act pay the same percentage of their wages. The taxes paid in the first instance by the employers will tend to be shifted either to the workers or to the consumers, and will be regressive in operation. It might be better, from the point of view of equity, to support the old-age and unemployment projects out of general revenues, but the tax consciousness promoted by the payroll taxes may be desirable.

The Estate and Gift Taxes. The combination of estate and gift taxes produced only 1.3 per cent of total federal revenues in 1945. The estate tax is applied to estates as a whole rather than to shares received by individual heirs. The first $60,000 of an estate is exempt, and the tax rates in 1945 ranged from 3 per cent on the parts of estates which barely exceeded the exemption to 77 per cent on the parts of estates in excess of $10,000,000. Credit was given against the federal estate tax for 80 per cent of any amount which was paid in taxes under state inheritance tax laws. The gift tax, which is necessary to prevent the evasion of the estate tax through transfers of wealth between living individuals, had rates which were three fourths as great as the estate tax rates.

The estate and gift taxes can be relied upon to produce a fair amount of revenue, but they are not a good source of increased revenue if the increase must be made available suddenly. The estate tax unquestionably accords with the principle of ability to pay. Inherited wealth is purely a surplus return to the heir, and is entirely unearned by him. The greater the amount that society permits to be passed on through inheritance, the greater is the share which society may legitimately claim from an estate. The incidence of the tax is clear. Its burden rests wholly upon the receivers of an estate and cannot be shifted. The tax is certain as to amount, and as to time and manner of payment. The tax would be objectionable if it interfered with the ability of individuals to provide for their heirs and dependents to such an extent as to lessen greatly the incentives and efficiency of the individuals and to slow down unduly the process of saving and capital formation. It does not seem, however, that the present tax is likely to have these effects. Again, if a reasonable period of time in which to pay the tax were not given, heirs might be forced to liquidate real estate, factories, machinery, and securities at a considerable loss in order to pay the tax. In such a case 40 or 50 per cent of the value of an estate might be lost in paying a tax of much smaller nominal rate, but this objection does not apply very strongly to the actual federal estate tax.

Excise Taxes. Excise taxes of various kinds produced 12.5 per cent of the total revenues of the federal government in 1945. Excise taxes are taxes on economic goods. In some cases, producers are required to buy and affix stamps to the articles they make and sell. Again, the producers may be required merely to pay the government a certain amount per unit of product produced and sold. Some of the taxes are specific, as for example a tax of, say, 8 cents per package of twenty cigarettes; while others are ad valorem taxes, such as a tax of 10 per cent on the factory price of automobiles.

While excise taxes are usually collected from the producer in the first instance, their burden is ordinarily borne by the ultimate consumer when the taxed goods are made under competitive conditions. The tax is shifted through increases in the prices of the taxed articles, for excise taxes are costs of production to the producers of the goods and must be covered by price if production is to continue. However, the process of shifting is not the simple one of adding the taxes to the old prices. According to the Law of Demand, an attempt to raise the price of a taxed article will decrease the volume of sales. The extent to which an increased price will curb sales depends upon the elasticity of demand within the relevant range of prices, but some decrease in sales is inevitable if the total demand for the good remains the same as before. A decreased volume of sales results in a decreased output of the good, and this change of output is likely to cause a change in the manufacturing costs of producing the good in the intermedi-

ate period. Since the new manufacturing cost may be either higher or lower than before, depending upon whether the former output was greater than, equal to, or less than normal capacity, the price of the taxed article may increase by either more or less than the amount of the tax, in the process of shifting the tax to consumers. The effect upon price also depends to some extent upon whether the tax is specific or ad valorem.

Under conditions of monopoly, partial monopoly, or monopolistic competition, the shifting of excise taxes is less certain than under competition. The enterpriser under such market conditions does not care who pays the excise taxes. His only concern is to set output and price at figures which will bring him the greatest possible total net profit from the production and sale of his good. In some cases, his former output and price may have been so much more profitable than any other combination of output and price that he will be better off in terms of total net profit to leave both unchanged and bear the tax himself, instead of attempting to shift it to consumers by raising price and reducing output. In other cases a new combination of output and price would doubtless be more profitable than the old one, after the imposition of the tax, and as a consequence the enterpriser would raise price and decrease output.

In general, the burden of taxes on commodities and other economic goods, no matter where it is first placed, tends finally to fall as a whole or in large part upon the consumers. Because of this fact, these taxes, judged by themselves, do great violence to our principle of ability to pay. They are not progressive, nor are they even proportional to income. People with large incomes pay these taxes in greater absolute amounts than people with small incomes, but the percentage of total income spent for economic goods subject to excise taxes tends to decline as a person's total income increases. Therefore, these taxes take away a smaller *percentage* of a large than of a small income, and are regressive in operation. This does not mean that they should never be imposed, for their bad effects might be quite offset by other taxes in the system. Excise taxes have been good revenue producers in the past, and have been considered a rather elastic element in the tax system. Their convenience in collection and payment is well known, for they "get the feathers with a minimum of squawking." They are often included in the prices of articles in such a way that most consumers are unaware that they are paying them.

Customs Duties. Customs duties, or duties on imports, produced less than 1 per cent of total federal revenues in 1945. These taxes have lost importance rapidly, for at one time they produced almost all the revenue of the federal government. With the United States operating under a high protective tariff, revenue from import duties is largely incidental. The main purpose of the tariff is to protect American industries by excluding foreign products from our markets, and a protective tariff which is completely

successful in this respect produces no revenue. Import duties paid on commodities from abroad ordinarily will have the same incidence as taxes on the production of domestic commodities; that is, they will fall on the final consumers. Hence these taxes may be criticized in much the same terms as those applied to excise taxes above, except that the fiscal adequacy of customs duties is more seriously open to question than that of excise taxes.

The Federal Tax System as a Whole. In some respects, it is difficult to summarize our study of the federal tax system. Some of its taxes are direct, others are indirect. Some are convenient to pay and economical to collect, others are not. Some are progressive, others are regressive. Certain generalizations may be made, however. The federal tax system as a whole meets fairly well the test of fiscal adequacy. In depression or in war years, of course, it often fails to provide enough revenue to cover all the expenditures which must be made, but it is difficult to imagine a tax system that would be perfectly elastic and adequate for all emergencies. In 1945, about 75 per cent of total federal revenue was derived from taxes which were at least nominally progressive. In the case of taxes on corporate incomes, however, there is no certainty that the progressive rates actually levied are converted into progressive rates on the incomes of individual owners of corporations.

The federal tax system as it existed in 1945 may have been necessary or even desirable for purposes of war finance, but it would be open to serious objection in peacetime because of its probable effects of restricting production and employment, and discouraging business ownership and the foundation of new enterprises. The heavy excess profits taxes used in 1945 would have no place in a peacetime tax system, for they would deprive many firms of any incentive to do more business than had been done in the base period and would place grave difficulties in the way of new firms which needed to increase their capital out of earnings. The excess profits taxes have been repealed, of course, but a number of further changes in the federal tax system would be desirable. The combined normal tax and surtax on corporate income should be reduced to a greater extent than was provided by the Revenue Act of 1945, the exemption should be raised, and more liberal provision for carrying losses forward as an offset to current income should be introduced. In the case of the personal income tax, it would be desirable to apply differential tax rates in order that personal incomes derived from self-employment (operating enterprises) and from corporate dividends might be taxed much less heavily than personal incomes derived from salaries or interest. And personal incomes invested in new plant and equipment, or invested in corporate stocks for the same purpose, might well be taxed at less than the regular surtax rates. The changes would be intended to stimulate production, employment, business ownership, and the founding of

new enterprises while producing a high level of revenue for the federal government.

State and Local Tax Systems. State governments in the United States in 1944 derived 32.6 per cent of their total revenues of $4,499,000,000 from sales and use taxes, 15.3 per cent from taxes on gasoline and other motor fuels, 17.0 per cent from income taxes, 8.9 per cent from motor vehicle licenses, 9.1 per cent from nontax sources, 6.4 per cent from business licenses, 5.4 per cent from property taxes, 2.5 per cent from inheritance and gift taxes, and 2.8 per cent from miscellaneous taxes.[19] In 1943, cities with a population of 25,000 or over obtained 66.3 per cent of their total revenues from property taxes, 16.9 per cent as aid from other governmental units, 9.2 per cent from miscellaneous taxes, and 7.6 per cent from public service enterprises and other nontax sources.[20]

The income, inheritance, and payroll taxes levied by state governments are not very different in their general characteristics from those levied by the federal government, and state taxes on gasoline and other motor fuels are similar to federal excise taxes. Motor vehicle licenses are an obvious source of state revenue. Most, if not all, states require the payment of a registration fee annually on all motor vehicles, while many collect additional fees for operators' licenses. The states collect fees in similar fashion for licenses to carry on certain businesses or to follow certain professions. For example, one must obtain a license and pay a fee to sell cigarettes, to operate a tavern, or to work as a barber. While these various fees are not taxes in the strict sense of the term, they have about the same effect as taxes on consumption or taxes on the operation of businesses which are able to shift the burden to the consumers. Only two types of taxes used by state and city governments—the sales tax and the general property tax—require detailed comment here.

Sales and Use Taxes. The sales tax has come into prominence in recent years as a source of revenue for state governments. The tax has been applied in a large number of states, is based on retail sales, and usually runs from 1 to 3 per cent on the value of the articles sold. Use taxes are usually intended to prevent people from escaping sales taxes—just as gift taxes are designed to keep them from avoiding inheritance taxes—and have been adopted by most states which use sales taxes. Use taxes are levied as a charge for the privilege of storing, using, or consuming within a state any goods which have been purchased outside the state. A person who lives in a state that has a sales tax, and who purchases goods outside the state to avoid the sales tax, becomes subject to the use tax when he brings the merchandise home.

The sales tax—and this is true of the use tax as well—is decidedly

[19] National Industrial Conference Board, *op. cit.*, p. 123.
[20] *Ibid.*, p. 135.

defective from the point of view of our principles of taxation. It is regressive in operation and does not conform to the principle of ability to pay. People with large incomes spend a smaller percentage of their incomes on the retail purchases subject to the tax than do poorer people, so that the tax takes a higher percentage of small than of large incomes. The sales tax is not so well received by the payers as some regressive taxes, for most of the sales tax laws require separate charging of the tax to retail purchasers, in order to promote the shifting of the tax to consumers and to make them conscious of the fact that they are paying it. The sales tax is rather costly to collect and not at all simple to administer.

The General Property Tax. Most of the revenues derived by state and local governments from property taxes come from the general property tax, which is a tax on property considered as a homogeneous whole and is sometimes called the "uniform rule" of taxation. This means that the rate of the tax is to be uniform throughout the taxing district and for any amount of property. The tax is based upon the valuation or assessment of property in terms of money. These valuations are estimated by assessors at specified times, and the tax is applied at a certain rate, ordinarily upon the total valuation of the property of the taxpayer in question. Boards of review are often created for the purpose of correcting inequalities and obtaining a uniform basis of assessment.

The general property tax is based upon the assumption that ability to pay is adequately represented by the ownership of "general property." The tax is defective in several respects. In the first place, it is based on a mistaken notion as to the nature of property. Properly speaking, property is an institution which guarantees to the individual the right to use and control, to receive benefit from the ownership of, to exclude others from the use of, and to pass on to others, at the time of death or before, whatever economic goods he may acquire. What is called "property" under this tax is in reality made up of two classes of things—wealth and claims upon wealth. Now when certain items of wealth and claims on these same items are both regarded as general property subject to tax, it is clear that double taxation results. Thus the corporation is taxed upon certain items of wealth, such as buildings and machinery, and the stockholder is taxed upon his shares of stock, which are claims upon these articles of wealth. This is double taxation in the worst sense because it is largely if not entirely unintentional, and because an item of wealth and a claim upon that wealth are both taxed at the full rate charged other items of wealth which are not represented by similar claims.

In the second place, it is assumed, at least by inference, that nothing other than "property" represents ability to pay taxes. This is clearly untrue at present, for many persons have very large incomes derived from personal services, but possess little wealth that is reached by the general property tax.

A third and important defect of the general property tax is that it is grossly inequitable. The assessment is made by assessors who are ordinarily dependent for their positions upon the votes of the people whose "property" is to be assessed. They are for the most part untrained and inexperienced, and do their work in a relatively short time. The result is incquitable valuation and taxation. Much intangible property (that is, claims upon wealth) escapes taxation altogether, while real and tangible property is assessed in a most discriminatory fashion. As expenditures have increased on the part of the governmental units dependent upon this tax, it has been found necessary to increase the rate at which the tax is applied. This has given the owners of intangibles an increased incentive to evade assessment, which means a smaller amount of "property" of this kind assessed and a still higher tax rate, which in turn stimulates further evasion, and so on. Evasion of the tax by the owners of intangibles is rather easy because the correct assessment of these items depends so largely upon the cooperation of the taxpayers themselves.

Although in theory the tax is based upon proportion (that is, the same rate being intended to apply, regardless of the amount of property an individual has), it seems altogether fair to say that the tax has been regressive in operation. The owners of great wealth are able to consolidate much of their holdings in forms which may escape the tax, and it is well known that real and tangible personal property is progressively underassessed as the amount possessed by an individual increases. Thus, the rate actually paid upon "general property" tends to be lower, the greater the amount of property possessed by the individual. In addition, it becomes more and more difficult as time goes on to insure the fiscal adequacy of the tax, for it is not easy to adapt the tax to increasing fiscal needs. It is decidedly inelastic.

It is not feasible to discuss the incidence of the general property tax as a whole. It must be considered as it falls upon owners of different kinds of wealth. In so far as the general property tax is a tax on land, the tendency is strong for the burden to rest finally upon the owners of the land, regardless of whether they pay the tax originally. Land is not a commodity produced at the will of men, but for all practical purposes is fixed and nonextensible in amount. The rent of land and consequently its value are determined by the conditions of demand for and supply of land. There is nothing about a tax on land which will make its user able to pay a higher rent for it, nor will its supply be affected in any way by such a tax as is ordinarily applied. Therefore the tendency is for the owner of land to bear the burden of any tax placed on it.

The incidence of the tax tends to be quite different when it falls on buildings, rather than on land itself. A tax on buildings tends to fall upon the tenant rather than on the owner, unless these two happen to be the same person. Buildings are simply one form of the investment of capital,

and if a tax is imposed which falls temporarily upon the owners and makes the return from this investment smaller than that which can be obtained in other fields, the tax tends to be shifted. The process of shifting is a long-run process, and comes about through an exodus of capital funds from the taxed form of investment, which enables the capital that remains in this field to receive as high a return as it received before the imposition of the tax.

A tax on buildings, therefore, seems to be merely one case of the taxation of capital. In so far as a tax falls upon some forms of investment and not upon others, or upon some more heavily than upon others, there is a tendency for the tax to be shifted through the process outlined in the preceding paragraph. In so far as a tax affects all capital alike, its shifting or nonshifting depends upon whether, in the long run, the tax is sufficiently high to operate as a check upon saving and capital accumulation. This possibility has already been discussed in other connections. To the extent to which the general property tax is imposed on articles of wealth which will be used not as capital in further production but merely for personal consumption, the prospect of shifting the tax is practically nil, because such articles do not ordinarily enter into later price transactions.

State and Local Tax Systems as a Whole. State and local governments, like the federal government, have often had great difficulty in making ends meet under their present revenue systems, and have become involved in borrowing operations. Both state and local governments derive a very large part of their revenue from taxes which do not conform to the principle of ability to pay. This means that the revenue systems in question are highly regressive in operation, and that the burden of state and local taxation falls relatively more heavily upon the poor than upon the rich. On the other hand, as these governmental units turn to the taxation of incomes and inheritances, they invade fields that are already heavily taxed by the federal government. Some people think, therefore, that it might be desirable to have only one tax system in the United States, with all taxes collected by the federal government and the revenue divided among the various governmental units.

GOVERNMENTAL BORROWING AND THE PUBLIC DEBT

The Growth of the Public Debt. The expenditures for carrying on governmental activities should ordinarily be made out of current revenues from taxation and other sources. In recent years, however, all of our governmental units have often found it impossible to secure sufficient current revenue to meet their expenditures and have depended upon public borrowing to meet the deficit. Although the debts of state and local governments have been increasing, we shall concentrate our attention upon the public debt of the federal government. This debt reached $26,597,000,000

after World War I, but by 1930 it had been reduced to $16,026,000,000.[21] Since 1930, the federal government has had a large deficit every year and its debt has been mounting rapidly. It reached $45,890,000,000 in 1939, after nine years of deficit financing of depression activities. Then, with heavy national defense and war expenditures added to an already unbalanced budget, the debt grew rapidly to $76,991,000,000 in 1942, $140,796,000,000 in 1943, $202,626,000,000 in 1944, and $258,682,000,000 in 1945, as shown in Table 26.

Table 26: Expenditures, Revenues, Deficits, and Public Debt of the Federal Government, 1931–1945*

(*In millions*)

Fiscal Year	Expenditures (during fiscal year)	Revenues (during fiscal year)	Deficit (for fiscal year)	Public Debt (end of fiscal year)
1931	$ 3,560	$ 3,008	$ 552	$ 16,801
1932	4,640	1,905	2,735	19,487
1933	4,611	2,009	2,602	22,539
1934	7,362	3,052	4,310	27,734
1935	9,738	3,724	6,014	32,824
1936	9,098	4,062	5,036	38,497
1937	7,754	4,973	2,781	41,089
1938	6,977	5,609	1,368	42,018
1939	8,361	4,990	3,371	45,890
1940	8,824	5,175	3,649	48,496
1941	12,774	7,727	5,047	55,332
1942	33,980	12,560	21,420	76,991
1943	79,604	21,982	57,623	140,796
1944	95,129	43,608	51,521	202,626
1945	100,405	46,356	54,049	258,682

* Sources: National Industrial Conference Board, *op. cit.*, p. 108, and Treasury Department release of August 1, 1945.

Inflationary Dangers of Public Borrowing and an Increasing Public Debt. There can be no sound objection to governmental borrowing to meet an emergency, when revenues fall short. If a national emergency requires expenditures in excess of collectible revenues, the federal government would be foolish not to permit its budget to become unbalanced. When a man needs a surgical operation, he does not hesitate to call in the surgeon merely because the cost would unbalance his budget for that year. On the other hand, the fact that the use of sulfa drugs may be indicated when a man has pneumonia does not mean that these drugs would prove beneficial as a steady article of diet. That is, continuing deficits of many billions of dollars per year may lead to the destruction of the government's credit. If the public debt grows so large that full payment of interest and repayment of principal become improbable, the government may be unable

[21] Daily Statement of the United States Treasury, June 30, 1941.

to continue borrowing from its citizens. If it is then unable or unwilling to decrease expenditures or increase ordinary revenues, it may resort to printing paper money with which to pay its bills. These inflationary tactics on the part of a government usually lead to an economic breakdown from which a country may not completely recover for many years.

In the period of World War II our federal government did not resort to the printing of paper money in order to pay its bills, but it did something not very different. Desiring to make total expenditures far in excess of the sums which it was thought possible or feasible to collect in taxes or to obtain from direct sales of bonds to the citizens, the government resorted to sales of bonds to the banks in order to obtain the necessary funds. The banks paid for these bonds by creating demand deposits for the government to spend. In this way the government obtained large sums of purchasing power without reducing the funds available for civilian spending, and the total of government and civilian purchasing power became much larger than necessary to take off the market at stable prices all the commodities and services which the economy could produce. The situation which was created was about as inflationary as that which would have resulted from the printing of paper money to finance governmental expenditures, except that the citizens apparently are not driven into as great a panic by the creation of demand deposits as by the printing of paper money. A tremendous inflation of prices was avoided during the war period by resorting to direct price control, production control, rationing, and other devices, but this only postponed the problem and did not eliminate it.

The Future Burden of the Public Debt. Another complaint about our rapidly growing public debt holds that we are arranging to pass on to our children and grandchildren a staggering burden of debt which they will have to pay, to their own great detriment. This contention is difficult to evaluate. Clearly, in the sense of *real income,* future generations may not suffer from a large public debt incurred now. If the government borrows to buy wheat to feed its starving citizens, the wheat is taken not from the crops of thirty or fifty years hence, but from present supplies. If the government uses borrowed funds to induce farmers to plow cotton under, current crops are reduced rather than those which our children will harvest. When the government spends billions of borrowed dollars for war goods, it causes a shortage of consumers' goods here and now rather than in the more or less indefinite future.

What really happens when the government borrows directly from the citizens is that the citizens are induced to turn over to it a part of their current money income. With these funds, the government is able to secure a larger share of our *present national real income* than it could otherwise obtain, and private individuals must get along with a smaller share than they would otherwise have. In the sense of real income, then, the cost of

public borrowing is borne now in the form of a smaller real income for private purposes than would otherwise be available. And the effect may be much the same when the government borrows from the banks, instead of directly from the citizens, provided ways can be found of sterilizing part of the large money incomes which will remain in the hands of the citizens.

From the *financial* point of view, it is certainly true that the government will have to collect taxes in the future in order to pay interest on its obligations, and possibly to pay part of the principal, but at the same time these amounts will be paid to the owners of the government bonds which represent our public debt. If the people who own the bonds also pay the taxes, even private individuals may "break even" on the process. But whether certain *individuals* gain, lose, or break even, it is clear that the *nation as a whole,* in paying off the public debt, merely transfers money from one pocket to another so long as the entire debt is held within the country. Such transfers, it is said, should not be very harmful.

On the basis of these and other arguments, some writers are inclined to scoff at the idea that our large public debt will give us trouble in the future. The size of the debt is a matter of small importance, since "we owe it to ourselves." While individual bondholders will need to be paid off from time to time, the principal sum of the debt as a whole will never have to be paid and our only concern should be the size of the interest payments in relation to our national income. Indeed, some writers seem to consider a continually growing public debt as a desirable thing, on the ground that large governmental expenditures in excess of current revenues will ensure annually the existence of a total sum of purchasing power in the country large enough to take off the market all the commodities and services which the economy can produce at full employment. If we can keep the national income constantly growing in this way, even a total of interest payments which steadily increases in absolute size will not be a very severe burden.

In my opinion, however, a lighthearted attitude toward the future burden of the public debt is far from completely justified. When taxes are collected to pay interest on the debt, they may be so large in amount or collected in such ways as to hamper and restrain production and employment and impair the effectiveness of our productive facilities. Moreover, increases in the total of fixed income payments in the economy, such as interest payments on government bonds, tend to concentrate the risks of the system on those relatively few incomes which are drawn from the operation of business enterprises and the investment of venture capital. It has been estimated that the annual interest on a federal debt of 300 billion dollars would exceed the total of interest on private debt plus rents in a full-employment national income, and would equal one fourth of total business income (dividends plus income of unincorporated businesses plus agri-

cultural income).[22] Again, with the banks loaded to the gunwales with government bonds, the government would certainly be much concerned about any general rise in interest rates which would cause government bonds, with their low fixed rates of interest, to depreciate sharply in value. Finally, the political implications of having a large class of persons dependent upon interest on government bonds for a considerable part of their income may be anything but favorable.

If we consider interest payments (and possible payments of principal) on the federal debt in combination with the possible cost of other federal projects now contemplated, the outlook even for the continued operation of a capitalistic system in the United States is rather unfavorable. Interest at 2 per cent on a national debt of 300 billion dollars would amount to 6 billion dollars per year, and this would be merely the cost of carrying the debt. If any serious attempt were made to pay off the debt gradually, further large expenditures would be necessary. If the debt were not to be reduced in this fashion, it would have to be refinanced and probably at higher rates of interest, which would have much the same effect on federal expenditures. Some people who are not much concerned about the size of the national debt seem to think that there is a limitless market for government bonds in this country at an interest rate of about 2 per cent, but we may doubt whether this would be true in peacetime and in a free market.

According to most pronouncements on the subject, our military and naval establishments and the armed forces in general will have to be maintained at a high level for some years to come, in order to protect ourselves from unknown or unmentionable dangers and to be able to play our part in international organizations for keeping the peace. The cost will surely run into several billions of dollars annually. We shall have other large expenditures for pensions, hospitalization, medical care, education, training, rehabilitation, and other projects undertaken for the benefit of veterans of World War II. Then we must add in some billions of dollars for the operation of ordinary departments of government; other billions of dollars for a system of social security, popularly known as the "cradle-to-the-grave" variety, which will cover many more people, provide against more risks, and furnish much larger benefits than the present system; and possibly still other billions of dollars for a federal guarantee of employment if the blessed condition of full employment does not result naturally.

Just how large the annual expenditures of the federal government would be on the basis of all these projects is anybody's guess, but most of the available guesses or theoretical models call for federal expenditures which are far larger than those to which we have been accustomed in peacetime. Some people are thinking in terms of only 18 or 20 billion dollars

[22] *Financing American Prosperity.* New York: Twentieth Century Fund, 1945, pp. 138–139.

per year. Others see a possibility that federal expenditures may run to 30, 35, or 40 billions annually. However great the annual amount might be, the large expenditures of the federal government in the postwar period would have to be made out of revenues from taxation if we did not wish to add steadily to the already overgrown federal debt. This would call for a continuation of very high levels of taxation, and possibly of levels not far removed from that which prevailed during the latter years of World War II.

The point in all this is that our economic system probably could not operate in capitalistic fashion in peacetime if the government found it necessary to take a very large part of the earnings or "profits" of business enterprises of all types through taxation, while leaving these same enterprises to bear almost unassisted any losses which they might encounter. And certainly such a situation would be most unfavorable for the expansion of the economy and the foundation of new enterprises on the basis of private capital. Various reforms of the federal tax system, which would make taxation bear less heavily than at present on the incomes which individuals and firms derive from employment-creating activities, might go far toward making heavy taxation, if not compatible, at least less inconsistent with the maintenance of a high level of employment and production in the private sector of the economy. However, the view that Congress is sure to enact such reforms merely because they would be desirable seems decidedly overoptimistic. On the whole, then, the size of the public debt can hardly be regarded as a matter of indifference.

QUESTIONS AND PROBLEMS

1. Why is it desirable to have certain services provided by governments rather than by private agencies?

2. What has been the trend in recent years in the volume of expenditures made by the federal government? Explain the nature of the principal items of expenditure.

3. "The chief functions and expenditures of state and local governments are about the same as those of the federal government." Show whether you agree.

4. "The growth of population has been the chief cause of increases in public expenditures in recent decades." Do you agree? Explain.

5. "The great increase in governmental expenditures and services in the great depression following 1929 was the result of a new theory in regard to the functions of government." Explain.

6. What is the general basis for reaching a decision as to the "proper" scope of governmental activities and expenditures? Explain.

7. "It is important to consider the indirect as well as the direct effects of governmental expenditures and collections of revenue." Explain.

8. Evaluate the depression activities and expenditures of the federal government.

9. What must we consider in evaluating federal activities and expenditures which are proposed for the postwar period? Explain.

10. "A sound general system of taxation should be based on a number of principles." Explain.

11. Does taxation according to ability to pay require regressive, proportional, or progressive taxation? Explain.

12. Distinguish between the so-called direct and indirect taxes.

13. What are the principal characteristics of the personal income tax as applied by the federal government? Is it a direct or indirect tax? Explain.

14. "The federal personal income tax is a well-nigh perfect tax." Show whether you agree.

15. "The federal tax on corporate income does not conform to the principle of ability to pay and results in double taxation." Do you agree? Explain.

16. "The capital stock tax and declared value excess profits tax operate in complementary fashion." Explain.

17. "The main federal excess profits tax, as applied during the war, should be made a permanent part of the federal revenue system." Do you agree? Explain.

18. How does the federal government tax estates? Explain.

19. What are excise taxes and who finally bears their burden? Explain.

20. Evaluate the federal tax system as a whole and make suggestions for its improvement.

21. What is the sales tax and how is it used? Is it a desirable tax? Explain.

22. "The general property tax is defective in several respects." Explain.

23. What is the incidence of the general property tax? Explain.

24. "There can be no sound objection to financing governmental expenditures through public borrowing." Do you agree? Explain.

25. "Heavy public borrowing is likely to lead to inflation." Discuss.

26. "The public debt of the federal government is certain to be burdenless in the future because we owe it to ourselves." Show whether you agree.

27. To what extent is it true that the burden of the public debt cannot be transferred to the citizens of future years? Explain.

28. "A lighthearted attitude toward the future burden of our large federal debt is far from completely justified." Do you agree? Explain.

REFERENCES FOR FURTHER READING

Buehler, A. G., *Public Finance.* New York: McGraw-Hill Book Company, Inc., 1940, Chaps. 1–6, 15–34.

Groves, H. M., *Financing Government.* New York: Henry Holt and Company, 1939, Chaps. 2–18, 23–27, 30–33.

Hunter, M. H., and Allen, H. K., *Principles of Public Finance.* New York: Harper & Brothers, 1940, Chaps. 1–5, 9–30.

Lutz, H. L., *Public Finance.* New York: D. Appleton-Century Company, Inc., 1936, Chaps. 2–5, 8, 15–33.

Mering, Otto von, *The Shifting and Incidence of Taxation.* Philadelphia: The Blakiston Company, 1942, Chaps. 1–16.

Schultz, W. J., *American Public Finance.* New York: Prentice-Hall, Inc., 1938, Chaps. 1–4, 10–23, 26–28.

XXIX

Economic Instability

In 1929 the United States produced a national income amounting to $83,300,000,000, but in 1932 it was only $40,000,000,000. Over the same period of time the number of persons employed in the United States decreased from 35,501,000 to 26,195,000, the index number of industrial production (1923–1925 = 100) fell from 119 to 64, and the index number of wholesale prices compiled by the Bureau of Labor Statistics declined from 95.3 to 64.8. The gross farm income of the country went down from $12,791,000,000 to $5,562,000,000; our international trade declined from $9,640,000,000 to $2,934,000,000 in terms of merchandise; and the ton-miles of freight carried by our railroads fell from 450,189,000,000 to 255,309,000,000. About the only thing that increased, apart from the number of persons unemployed, was the number of industrial and commercial failures, which went from 22,909 to 31,822. In the financial world, the loans and discounts of all active banks fell from $41,433,100,000 to $28,089,900,000; the deposits of these banks from $53,244,700,000 to $41,126,000,000; securities issued from $11,592,200,000 to $1,730,300,000; stock prices from an average of $311.90 to one of $80.88 per share; brokers' loans from $8,549,000,000 to $244,000,000; and stock sales on the New York Stock Exchange from 1,125,000,000 to 425,000,000 shares.[1]

What happened to economic activity in the United States between 1929 and 1932? This is a question which anyone can answer. Economic activity in the country merely passed from a state of great prosperity to one of extreme depression. In view of the magnitude of the changes which occur in economic activity as between prosperity and depression, it is apparent that we would be overlooking an important department of economic experiences if we concentrated attention entirely on the long-run, equilibrium aspects of prices, income distribution, production, domestic exchange, and international trade. Therefore, in the present chapter and in the one which follows, we shall examine the subject of economic instability at considerable length.

[1] All these data are from the *Statistical Abstract of the United States*, 1939 and 1941.

THE EFFECTS OF BOOMS AND DEPRESSIONS

Changes in Price Levels. Periods of prosperity and depression have very severe effects on the economic welfare of the people of our economic system. In the first place, the general price level tends to rise in prosperity and fall in depression, and these changes in the price level benefit some persons and injure others. An increase in the general price level means a fall in the purchasing power of money. Any person whose money income remains fixed when the price level rises, or whose money income rises less rapidly than the price level, loses purchasing power and real income, while anyone whose money income rises more rapidly than the price level registers a gain in these respects. Conversely, since a fall in the general price level means an increase in the purchasing power of money, anyone whose money income remains the same, or falls less rapidly than the price level, gains in purchasing power and real income, while anyone whose money income declines more rapidly than the price level suffers a loss in these respects.

Thus debtors tend to be benefited in periods of prosperity. If their money incomes increase, it may become relatively easy for them to pay off their fixed debts, and in any case they will return to the creditors sums of money which will have less purchasing power than they had when they were borrowed. Business enterprisers are likely to gain as debtors, and they tend to gain in other ways as well. As we noted in Chapter XXI, the processes of indirect production require the business enterpriser to incur many costs of production some weeks or months before the final products are ready for sale and, if the price level is rising, the prices of these products when they are sold may have ascended to a level well above that which was expected when their production was undertaken. Moreover, quite apart from the roundaboutness of production, some costs of production, such as rents, wages, and interest payments, may either remain fixed for a time as the general price level rises or will rise less rapidly than the price level in general. As a result of these factors, periods of prosperity and rising prices are usually periods of great profits for business enterprisers. Any persons who depend upon dividends from common stocks for their money incomes should be classed with other business enterprisers in this connection.

Creditors tend to lose in periods of prosperity and rising prices, as they are paid back given sums of money which have less purchasing power than when the loans were made. Persons living on fixed money incomes from pensions, annuities, or bonds also experience a loss of purchasing power and real income in such periods. Salaried workers and wage earners are in the same boat but do not usually get quite as wet. Their wages and salaries are likely to be raised in periods of prosperity and rising prices, but they usually do not increase as fast as the general price level. However,

workers as a group may be relatively well off because industry and trade are expanding and employment exists for almost all persons who are able and willing to work.

In periods of depression and falling prices, the tables are turned to some extent. Debtors who incurred obligations when prices were high may find it very difficult, if not impossible, to pay their debts, and, if they are successful in meeting their obligations, will still lose in a sense by returning to their creditors sums of money whose purchasing power is greater than that of the funds which they formerly borrowed. Business enterprisers lose as debtors and in other ways. The roundaboutness of production and the lagging of certain of their costs behind the general price-level movement, which operate to bring them profits in periods of prosperity, furnish them with losses to the extent that they keep on operating during periods of depression and falling prices. In similar fashion, persons who depend upon dividends from common (or even preferred) stocks for their money incomes are likely to find that these incomes decline faster than the general price level, if indeed they do not disappear altogether.

Persons who live on fixed money incomes from pensions, annuities, bonds, or other sources tend to gain in purchasing power and real income in periods of depression and falling prices if their money incomes really remain fixed. All too often, however, pensions will be reduced or corporations will be unable to pay the interest on their bonds, and prospective gains for these people will be converted into losses. Salaried workers and wage earners who are able to hold their jobs also tend to gain in terms of purchasing power and real income, as their salaries and wages fall less rapidly than prices in general in periods of depression, but this consideration means very little to the millions of prospective workers who may be entirely unable to find employment in such periods. The net conclusion must be that price-level changes in prosperity and depression are a matter of real concern in our economic system.

Changes in National Income. Business depressions may conceivably bring benefits to particular individuals, but from the point of view of the economy as a whole they bring loss and waste. In such periods the economy produces a national income annually which is far below that which it is capable of producing. From the data in Table 27 we can readily compute that, in the ten years after 1929, our total national income amounted to only $582,000,000,000, or about $251,000,000,000 less than the total income which would have been achieved if the 1929 national income had been produced in each year of the period. The total loss of income over the period would appear still greater if we took into account the fact that the economy should have been able not merely to maintain but actually to improve upon the 1929 national income in the later years. The loss of real income over the period was considerably less than the loss of money income, since prices were

Table 27: National Income of the United States, 1929–1945[2]
(*In millions*)

Year	National Income	Year	National Income
1929	$83,300	1938	$ 64,200
1930	68,900	1939	70,800
1931	54,500	1940	77,600
1932	40,000	1941	96,900
1933	42,300	1942	122,200
1934	49,500	1943	149,400
1935	55,700	1944	160,700
1936	64,900	1945	161,000
1937	71,500		

much higher in 1929 than in the depression years which followed, but it was very large nevertheless. And any such loss of real income seems particularly regrettable in view of the fact that the greatest income which could have been produced in these years would have fallen far short of satisfying the desires of all the people for commodities and services.

Changes in Employment. Corresponding to the changes in national income from prosperity to depression are other changes in the volume of employment for workers and other productive resources. In the eight years from 1922 through 1929, some 1,360,000 workers had been unemployed on the average.[3] This figure was probably not much above the irreducible minimum of unemployment, for some workers are sure to be in between jobs, moving from one place to another, or seasonally unemployed, regardless of the time at which a count is taken. Unemployment then increased rapidly to 12,478,000 in 1932 and 12,744,000 in 1933, and averaged 9,022,000 for the ten years from 1930 through 1939.[3] Unemployment of land, factories, stores, machinery, equipment, and other productive resources, though difficult to measure, inevitably accompanied this unemployment of labor. It is a pitiful sight to see large quantities of all sorts of productive resources wasting in idleness in the face of a great host of unsatisfied human wants.

Human Values. Any set of statistics on the effects of depression is bound to fall short of the mark, because statistics cannot well express the complete influence of depressions on human lives and values. The cold figures for business failures cannot tell us of the worry and anguish which

[2] These Department of Commerce estimates are for national income produced. Ideally, this income would equal the gross income of all producing units minus raw materials consumed by all producing units and minus depreciation and obsolescence of capital equipment. As actually measured, national income produced is found by adding estimates of corporate saving to national income paid out (wages, salaries, dividends, interest, rents, royalties, and entrepreneurial income withdrawn). The estimates omit such items as income from odd jobs, income from illegal pursuits, imputed income from owned homes and other durable consumers' goods, and extracommercial household services performed by housewives and other family members.

[3] *Financing American Prosperity.* New York: Twentieth Century Fund, 1945, pp. 10–11.

a man experiences as the result of the failure of a business which he has developed through hard work and sacrifice over the years. They cannot tell us what happens to him and his family as his source of income, savings, and even personal wealth are lost in the wreck. The mere statistics on unemployment do not say how a worker feels as he moves his family to poorer quarters, watches its members suffer from lack of food and fuel, and dissipates his meager savings to keep body and soul together. They do not picture the feeling of hopelessness and of not being wanted or needed which comes from the endless, fruitless search for work, the effect on the pride and morale of the worker when he is unable to obtain employment and must subsist on charity or relief, or the permanent decline in the worker's efficiency and productivity which may result from years of unemployment. Truly the effects of severe depression on the lives of the people may be little short of catastrophic.

BUSINESS CYCLES OR NO BUSINESS CYCLES

In view of the facts which have been presented, it is not surprising that the task of finding an explanation of booms and depressions has engaged the attention of a large number of economists (and engineers, newspaper writers, entomologists, steamboat captains, and almost everyone) for many years. And most of these analysts have been committed to the idea that booms and depressions are parts of a mechanism known as the "business cycle." According to this notion, the various phases or periods of the business cycle follow each other in regular and unending succession, and each cycle is similar to other cycles to such an extent that we may hope to find some single explanation for all of them.

The Period of Prosperity. Believers in the existence of a "business cycle" usually divide a cycle into four periods, called prosperity, recession, depression, and recovery. With a natural desire to consider pleasant things first, let us start with the usual description of the period of prosperity. In this period business is booming. Industrial production and other types of production are at their maximum and industries are straining to put enough goods on the market. Prices have been rising and are high in this period, while the earnings of business enterprises are pleasingly plump. There is relatively full employment of labor, as well as of other productive facilities, and wages are high, but the workers, conscious of their unusually strong bargaining power at such times, are quite likely to go on strike in the hope of further improving their status. The building of new plants and the purchase of new machinery and equipment are often undertaken in this period, while business and commercial failures are at a minimum.

There is a great expansion of credit in the period of prosperity. The loans and discounts of the banks increase rapidly and their deposits expand. Bank reserves decline and may approach the legal minimum requirements.

Large issues of securities are floated with the greatest of ease, and the investment banking business thrives. The velocity of circulation of money and credit, as well as their quantity, may increase; there is a large amount of speculation in commodities and securities; and the interest rate rises and becomes high as the period wears on. The psychological attitude which pervades the industrial, business, and financial system at this time is often one of almost unbounded optimism. If the prosperous period is really a boom, there may even be a feeling in the economic system that the country has reached a "new era" in which poverty will be abolished and business depressions will be a thing of the past.

The Period of Recession. At the moment we are not interested in the forces which allegedly cause economic activity in general to pass from one phase of the "cycle" to another. Since we are concentrating attention on the characteristics of the various "cycle periods," we turn at once to the period of recession which follows prosperity. In this period industrial and business activity is declining, and everyone is striving to achieve liquidity. Banks try to call in loans and cash in on investments, fearing that their depositors may ask for their deposits in cash. Businessmen, faced with an actual or prospective curtailment of their sources of credit and with large stocks of goods on hand which have piled up in the later stages of the preceding prosperous period, try to collect their business debts and may feel forced to sacrifice their goods at bargain prices. With prices falling, plenty of goods on hand, credit becoming restricted, and prospects of further profitable production virtually nil for the time being, industries cut their rates of operation, discharge employees or put them on part time, and curtail purchases of raw materials and supplies. Employment for workers therefore decreases and wages fall although workers sometimes try to resist the trend by going on strike. Many business firms and banks cannot succeed in putting themselves in a liquid position, and business and bank failures are on the increase.

Security prices fall sharply, and it is difficult to float new security issues, if, indeed, anyone wants to do so. The loans and discounts of the banks decline, as do their deposits, while reserves increase. The interest rate may remain high for a time, since many business and industrial firms are badly in need of funds to tide them over the emergency. Moreover, banks are hesitant about lending, but eventually the interest rate declines as the period moves along. There may be quite a little speculation on the "short" side of the commodities and securities markets. The psychological attitude in the business, industrial, and financial world is one of great apprehension. All participants in economic activity feel more or less like convicted prisoners before the bar of justice. That is, they know they have something unpleasant coming to them but they do not know just how bad it will be.

The Period of Depression. According to the traditional description, economic activity in the period of depression may be likened to a sailing vessel in the doldrums. Industrial and business activity is at low ebb at least until stocks of goods on hand are completely or well-nigh exhausted, and the prices of goods remain at a low level. Productive establishments are closed down or running at a fraction of capacity, and unemployment among workers is widespread. Wages, which lag behind falling prices in the early stages of the downward movement, eventually decline to a comparable level, and at the time workers are too weak in bargaining power to offer much resistance to the trend. Many business and industrial concerns that received mortal wounds in the recession period perish completely in the depression. Many banks fail also, finding that their fate was sealed by the events of the recession period even though they were able to survive that emergency.

Financial activity is also at a minimum in the depression. The loans and discounts of banks, and their deposits as well, have been considerably reduced. The rate of interest is low. The banks, since they are burdened with large quantities of idle funds, are very willing to extend credit and at a moderate rate of interest, but on the other hand, they are inclined to be most cautious and conservative in appraising prospective borrowers as credit risks. Speculative activity is exceedingly light in this period, and almost no security issues, except perhaps refunding issues, are offered for sale.

Business enterprisers cannot afford to produce many goods so long as the effective demand for these goods is small, and the total demand for goods can hardly be expected to increase so long as the various industries and businesses pay out only greatly restricted amounts of purchasing power for wages, raw materials, and other things. Sometimes it seems as if the economic system will never get away from its helpless "dead center" position, but it always does. The psychological attitude in the business, industrial, and financial world during depression is one of almost unbounded pessimism. Erstwhile captains of industry, business, and finance loudly proclaim that our country has reached the end of its career as an expanding and thriving economy, and that we can expect nothing better than continued hard times, world without end.

The Period of Recovery. As it is usually described, the period of recovery is the most attractive phase of the "business cycle," for it is a kind of cyclical springtime. Industrial activity is on the upgrade, employment is increasing, the money incomes of farmers, workers, and others are growing, and the prices of goods move upward from their depression lows. Wages and some other expenses of production lag behind to some extent as the prices of things in general rise, and the earnings of business enterprises improve. Workers, mindful of the recent lean times of unemployment and depression, are likely to be content with employment at almost any wages

for a time. However, as the recovery of business goes on, the workers begin to become aggressive again and, by forceful methods if necessary, seek once more to improve their economic lot. Nevertheless, business thrives and industrial and commercial failures become infrequent.

A similar rebirth of activity occurs in the financial areas of the economy. Borrowing at the banks picks up once more, a growth takes place in loans, discounts, and deposits, and idle reserves of funds decrease. Speculative activity in the commodities and securities markets increases and it becomes possible to sell new issues of securities. The interest rate may remain fairly stable during the early stages of recovery, but it tends to rise sooner or later as recovery goes on. The psychological attitude in the world of business, industry, and finance is one of hope and expectation. In the course of one complete revolution of the alleged "business cycle" we have now reached the point where we came in, and we must turn to other matters.

Overproduction Theories. The theories which have been developed to explain the "business cycle" are not only numerous but highly varied. They run all the way from theories which attempt to trace "business cycles" to such things as spots on the sun or eight-year cycles of the planet Venus to the self-generating theory which (some people think) holds that, although attempts to develop perpetual-motion machines have met with dismal failure in other fields, the economic system furnishes a first-rate example of such a contrivance. In the present connection it is not possible even to mention all of the theories, and brief comment will be made on only two of them.

Some of the most venerable theories of the "business cycle" hold that the explanation of these events lies in the periodic general overproduction of economic goods. With minor variations, these theories contend that the remarkable increases in the productive efficiency of our economic system over the years have brought this system to a point where production periodically outruns the demand for all sorts of goods. When such general overproduction occurs, the production of goods, like the hare in the fable, lies down and rests by the wayside until the tortoise-like demand for goods catches up, whereupon the productive system may begin to operate again.

Theories of this type overlook some rather fundamental economic principles. If human wants are really insatiable or are expanding without known limit, there is no reason to expect that, with the help of our scarce means of production, we can produce all kinds of goods in quantities which exceed the total desires of all consumers. Neither is it likely that the total demand for goods will be inadequate to take the total supply of goods off the market, for the total supply of goods *is* the total demand for goods. Whenever one produces a certain good for the market, he creates not only a supply of that thing but a demand for an equivalent amount of other

things. Thus the total demand for goods on the market consists of the total goods produced for the market. The total demand for and supply of goods on the market are not merely equal but identical.

From a somewhat different point of view, it is obvious that what the buyer of a good pays the seller receives. The total compensation or income for all the productive agents which cooperate to produce an article is measured by the price for which the article sells. More broadly, in producing a given total value of commodities and services for the market, the economy as a whole creates an equivalent amount of money income. In other words, national income equals national product, and it is impossible for the economy to create a volume of commodities and services whose aggregate value will exceed the aggregate means of payment of all the citizens.

The Self-generating Theory. The self-generating theory of the "business cycle" seems to hold that there is no single cause of cyclical fluctuations, that business cycles have no beginning and no end, and that each phase or period of the cycle contains the forces which operate inevitably to bring on the next phase. In the early stages of the upward movement, industrial activity is increasing and production is expanding. The demand for various kinds of goods increases and the price level rises. With prices rising, the operation of business enterprises becomes highly remunerative because of familiar factors such as the roundabout character of production, under which many costs must be incurred and be paid some weeks or months before selling prices are received for finished products, and the tendency of certain costs to lag behind the rising general price level.

With productive activities profitable, businessmen are anxious to expand operations to the utmost. They borrow heavily, take on more workers, expand productive facilities, and lay in large stocks of materials and supplies, or increased inventories of the goods which they carry as stock in trade. The increased demand of businessmen for capital goods stimulates the capital goods industries, and the costs of production paid out by these industries are spent and increase the demand for consumers' goods. Moreover, each increase in the operations of a consumers' goods industry tends to stimulate the demand for consumers' goods in general since the payments of the industry for costs of production are spent mostly for consumers' goods in the last analysis. Thus the whole upward movement in business is cumulative and grows on what it feeds on. Needless to say, a great expansion of credit is necessary to permit the sale of a growing quantity of all kinds of economic goods at continually increasing prices.

Why does such a prosperous period have to end? Obviously the expansion of productive operations can go on only to the point at which resources of production, human or otherwise, are fully employed. Further expansions of credit can only result in the rapid bidding up of the prices of the agents of production. As the prosperous period wears on, the costs

of production which formerly lagged behind the general price level begin to catch up. Bank credit becomes quite fully expanded and reserves are near the legal minimum requirements. It becomes increasingly difficult to sell a growing volume of production at higher and higher prices. Unwanted stocks of goods pile up on the hands of industrial enterprisers, and labor troubles complicate the situation. The demand for capital goods falls away and conditions in the capital goods industries begin to become depressed. Security prices begin to decline as confidence in the future is shaken. Storekeepers, merchants, and dealers stop stocking up with goods and attempt to cut down their holdings. Businessmen in general, remembering that prosperity has always given way to depression in the past, become apprehensive of the future. The demand for goods falls off everywhere except in retail markets. Eventually a few major industries decide that very stormy weather lies ahead and that they cannot safely maintain their existing rates of operation.

As these industries retrench they cut down production, lay off workers or put them on part time, curtail their purchases of raw materials, and so on. These activities cut down the incomes of workers, of farmers and of others who produce raw materials, of people who deal in materials, and others. The declines in these money incomes operate to decrease the demand for all sorts of economic goods, including those in retail markets, and industries in general begin to feel the change in business conditions. They all start to retrench and cut down expenses, incomes of various classes of persons are cut still further, the demand for various economic goods again declines, productive operations are cut still further, and so round and round. Moreover, industries producing consumers' goods feel repercussions from the thoroughly depressed conditions which exist by now in the capital goods industries. While all this is going on, the bankers are worried about their position and call in loans and restrict credit. Businessmen are driven to try to liquidate their unwieldy stocks of goods at sacrifice prices and to collect their own debts from other people, while their disinclination to maintain productive operations is increased. Many business enterprises and some banks are unable to meet their obligations and fall by the wayside—a fate which frightens the surviving businessmen and banks still more. Thus the downward movement of business, like the upward movement, is cumulative in character.

Eventually the storm is over and the survivors come out of their holes to survey the wreckage. The economic system settles down for a period of placid depression. Production and consumption continue on a restricted level until eventually things change for the better. Certain costs of production, which lagged behind falling prices during the early stages of the downward movement, finally fall as far as other prices. Debts have been liquidated. The capitalizations of business concerns are written down and

their fixed charges are reduced. Labor conditions are relatively peaceful and conditions in general are favorable for recovery. After a while the stocks of goods, which were carried over from prosperity in spite of the best efforts to sell them, become exhausted or are reduced as much as their holders desire. When this occurs, production has to be increased at least in some industries in order to fulfill even the relatively low depression demands for goods. These industries rehire some workers and begin to buy materials and supplies in increased amounts. The increased payments which these industries make on account of the costs of their increased production find their way directly or indirectly into the hands of persons who will spend them for consumers' goods. The demand for consumers' goods in general increases to some extent, more and more industries find it desirable to increase production, total payments to wage earners and raw material producers increase further, the demand for all sorts of goods increases another notch, and another cumulative upward movement of business is under way, to be followed in turn by another recession and depression.

In one respect, at least, the self-generating theory is a comfortable one to hold. "Business cycles" have no beginning and no end. No one in particular is responsible for them and nothing in particular is at fault, unless it be the type of economic system which we are trying to operate. From another point of view, the self-generating theory is depressing. "Business cycles" are an endless chain and there seems to be no point at which we could break it. In order to prevent a given depression, it would be necessary to prevent the prosperous period which preceded it, and that could be done only if we could prevent the previous depression, and so on.

The self-generating theory is often criticized for conveying too strong an impression as to the automatic character of the "business cycle." Such criticisms hold, in other words, that outside forces must intervene in the succession of booms and depressions from time to time, for they would run down if they depended entirely on internal forces. In this connection the most vulnerable part of the theory is the explanation of why recovery arises out of depression. Even if we grant that some slight increase in industrial activity must take place as stocks of goods become exhausted in a depression, what is to keep the recovery from stopping as soon as this slight increase has taken place? As a matter of fact, it has often been true in the past that some outside force or other has intervened in depression to help produce a real recovery, but we cannot expect such factors to be automatically present in a depression period.

Is There a "Business Cycle"? One fundamental objection to these theories of the "business cycle," and to others that could be mentioned, deals directly with the phenomenon which they seek to explain. In other words, is there any such thing as a "business cycle" which is amenable to explanation by these or any other theories? Periods of prosperity and

depression are easily recognizable, but are the entire upward and downward movements as a whole sufficiently repetitive to deserve the name of "cycles"? Are successive prosperities and depressions sufficiently similar with respect to duration, severity, and leading economic ingredients so that any single explanation can hope to account for them all?

In the analysis of actual periods of prosperity or periods of depression, a large amount of variation in duration and severity may be noted. The studies of the National Bureau of Economic Research show that the upward movements of economic activity since 1855 have varied in length from a low of nine months to a high of forty-six months. The downward movements of economic activity have been as short as eight months and as long as sixty-five months, while complete upward and downward movements have ranged in length between twenty-nine and ninety-nine months. In some cases the upward movement has been several times as long as the downward movement, while in other cases exactly the reverse has been true. Equally large variations occur in the amplitude or severity of the upward and downward movements in economic activity. Some prosperous periods are mild, peaceful affairs, and the depressions which follow are so moderate that they are scarcely noticed by most people. Other periods of prosperity are wild, feverish orgies, and the resulting depressions are so severe that they leave their marks on our economic system and its people for years to come.

General business, in which cycles are supposed to occur, is what is sometimes called a high-level abstraction. When we trace it down to a level at which we come in contact with concrete phenomena, we find that it is made up of hundreds of different phases of economic activity. Upward and downward movements in these different phases of economic activity show widely varying amplitudes and do not necessarily coincide even in the matter of timing. In a period of prosperity, some industries or fields of economic activity may share scarcely at all in the general prosperity of the times, while others prosper to an extreme extent. Similarly, in a period of depression, some industries may be in the most desperate straits, while others are running at nearly full capacity and are enjoying relative prosperity.

Each period of prosperity or depression is likely to contain some factors peculiar to itself. One period of prosperity may be greatly affected by a wartime demand for products. Another may center around the beginning and rapid development of one or two new industries, while a third may be influenced by the invention and adoption of new machines and productive processes in several industries. Other factors repeat themselves from prosperity to prosperity, but even these are likely to operate differently from one period of prosperity to another. Thus the general price level tends to rise in a prosperous period, but in one such period the increase in prices may

be severe while in another it may amount to comparatively little. In one period of prosperity the rise of the price level means that most prices are rising, while in another case most prices may remain rather stable while one or two groups of prices increase enough to push the general price level up. In similar fashion, bank credit is usually expanded greatly in prosperous periods, but in one prosperity the expansion may result largely from an increase in short-term commercial loans to business men (loans of the kind that involve commercial paper eligible for rediscount under the Federal Reserve System), while in another the expansion of bank credit may result almost entirely from an increase in other types of loans.

Some downward movements of economic activity are characterized by financial panics, but others are not. In some cases bank failures are concentrated in the early part of the downward movement, but in others most of them occur near the end of the movement when recovery is just around the corner. In some depressions, wage rates fall to very low levels while in others they are much better maintained. In some downward movements the prices of almost all commodities fall severely, while in other cases some prices remain relatively high while others bear the full brunt of the decline. On the basis of these considerations, we may conclude that periods of prosperity and depression do not make up highly repetitive cycles, and we may doubt whether any single theory is likely to be able to account for all periods of prosperity and depression.

UNDERINVESTMENT AND THE THEORY OF ECONOMIC MATURITY

Limiting our discussion now to the great depression which occurred in the years following 1929, we note that many people are confident that they have the explanation as to what happened in that period. The theory which they advance goes by many names, but it is most commonly known as the underinvestment theory.

The Significance of Aggregate Demand. While admitting that the economy as a whole, in producing any given total value of commodities and services, must create an equivalent amount of money income in the hands of the people, the holders of the underinvestment theory point out that another factor, known as aggregate demand, is most important in determining whether a high level of production and employment can be maintained in the economy. Disregarding governmental expenditures, aggregate demand consists of private expenditures for consumption plus private expenditures for investment in capital goods (including net foreign investment). How could this aggregate demand fall short of a sum adequate to take off the market all the commodities and services which the economy could produce at full employment? According to the common interpretation, this will happen if people take purchasing power out of the income stream by means of savings and do not return this purchasing power

to the income stream by means of investment—hence the name "underinvestment theory."

It should be apparent, however, that we could just as well say that aggregate demand will be deficient if people do not spend for consumption all the money income which they do not actually invest (directly or indirectly) in capital goods. From this point of view the explanation should undoubtedly be called an "underconsumption theory." Again, having a desire not to bias the discussion in either direction and noting that aggregate demand will be deficient if there is inadequate spending of money income either for consumption or for investment, or for both, we might call the explanation either an "oversaving theory" or an "underspending theory."

The Underinvestment Theory. In a prosperous period such as that of the late 1920's, a great burst of investment in capital goods usually occurs, and investment may go on at a rate which could not be maintained year after year without ruining the owners of existing capital goods through low rates of return. However, when investment slackens, a depression gets under way. With investment falling short of the sums which people save, aggregate demand becomes inadequate to take off the market all the commodities and services which the economic system is producing. Then the economic system must operate at a more restricted rate, but in doing so it creates a smaller flow of money income for the citizens. If the people would consume all of this limited money income except the very small sum needed for investment at such a time, a new balance might be struck. However, saving and consuming habits are relatively stable, saving even out of a restricted national income is likely to continue at an excessive rate, aggregate demand is again inadequate to take off the market all the commodities and services which the economy is producing, production and employment slump further, and so the depression goes on and on in the usual cumulative fashion.

Outlays for investment in capital goods are never a large part of total national money income, but apparently they are a very important part. For example, private outlays for investment fell from an annual rate of 17 billion dollars in 1929 to one of 2 billions in 1932.[4] This drop of 15 billion dollars in private investment was greater than any which had been experienced previously and it produced severe depression and unemployment in the capital goods industries. While this was going on, total expenditures for consumption declined almost twice as much, and the entire national income fell not by 15 billion dollars but by approximately 43 billions, or from a level of 83 billions to one of 40 billions. Thus a relatively small original amount of underinvestment, by making aggregate demand fall short of a sum adequate to maintain a high level of production, employ-

[4] *Financing American Prosperity*, p. 210.

ment, and national income, can apparently set in motion a cumulative spiral of depression in which production, employment, and national income will fall to fantastically low levels. Each drop in income and employment makes private investment in capital goods less and less attractive, and each fall in private investment, through its effect on aggregate demand, tends to lead to a lower level of income and employment.

The Remedy of Compensatory Spending. According to the underinvestment theorists, the remedy for the problem of underinvestment or deficient aggregate demand is relatively simple. When a depression threatens, measures to stimulate private investment would be largely ineffective, and measures to stimulate expenditures for consumption or to limit saving or soak up excessive savings would be slow working and uncertain in effect. Therefore the maintenance of an adequate level of aggregate demand at such a time appears to require public investment, or spending by the federal government in excess of its current revenues. Such spending by the government, over and above the sums which it receives from the current money incomes of the people, tends to compensate for the unnecessary and excessive savings of the people, and hence is known as compensatory spending.

If the government is ready to step into the breach with expenditures for a well-planned program of public works and developmental projects, a depression can be forestalled. Such governmental expenditures, if timely, would not have to be very large, because private investment would drop less in this situation than it would if left to its own devices, and expenditures for consumption would scarcely fall at all if the total level of investment were well maintained. Compensatory governmental spending to avert depressions would not necessarily add anything to the public debt in the long run, it is said, because the government could collect revenues in excess of its expenditures in periods of good business and thus retire the bonds issued to finance its compensatory spending at other times. Such a procedure might have the further beneficial effect of making periods of prosperity more moderate than they would otherwise be.

The Theory of the Mature Economy. Some underinvestment theorists attempt to provide a broader basis for their theory by advancing a further explanation sometimes known as the theory of economic maturity. Many years ago when our economic system was young, vigorous, and growing, the savings-investment problem, or problem of seeing to it that sums drawn from the income stream by means of savings are returned to this stream by means of investment, could hardly be of any great importance. Opportunities for investment in such a situation were almost boundless, and the risks involved in ambitious new business ventures were relatively small. The overexpansion of manufacturing plants or transportation facilities and the overbuilding of cities did little damage, for growth and expansion soon

canceled the errors. Men could safely undertake investment projects which were completely overoptimistic on the basis of existing needs. Now, however, our economic system is mature, senile, and decrepit. Because of the changes which have taken place in the system, there now tends to be a persistent shortage of investment outlets and a perennial tendency toward oversaving. Deficiencies of aggregate demand are chronic rather than occasional, and there is a continuing need for governmental stimulation.

What are the factors responsible for this change? One is the declining rate of population growth. When the population of the country was increasing rapidly, investment in and production of capital goods had to continue on a large scale both to provide a more abundant life for the existing population and to provide for the newcomers in the population. Needs for housing, transportation, and public utility facilities, for example, were very great. In modern times, however, the rate of population growth has fallen to a very low level, and our population threatens to become stationary in the not too distant future. Investment opportunities are rapidly becoming limited to more intensive investment for the benefit of the existing population. Estimating that population growth was responsible for about 60 per cent of the capital formation which occurred in the United States in the last half of the nineteenth century, the mature-economy theorists indicate that an important outlet for investment funds has almost disappeared.

Another important factor in limiting investment outlets has been the passing or disappearance of the frontier, so it is said. The rapid growth of population in this country in the past coincided with the settlement and exploitation of new territories. This development of virgin areas required heavy capital investments and furnished an important but nonrecurrent outlet for savings. When our own frontier had disappeared, we made more or less feeble efforts to use part of our savings in making investments in other backward areas of the world, but this outlet has not given us in the past, and does not promise to give us in the future, an adequate compensation for the passing of our own frontier.

The decreasing rate of population growth and the passing of the frontier combine to make us increasingly dependent upon technological changes for the existence of adequate investment outlets, but here again we are apparently doomed to disappointment. That is, there seems to be a dearth of great new industries, such as the railroad, electric power, and automobile industries in past times, which will give a lift to investment. When the automobile industry came into being and grew with extreme rapidity, for example, a very important investment outlet was created not only in the automobile industry itself but also in underlying and supplementary industries, such as those producing glass, rubber products, petroleum products, electrical appliances, steel, and cement. The effect of any one new major industry soon wears off, however, for the rate of new invest-

ment in it can be maintained only if the demand for its product continues to grow at an increasing rate. If the demand goes on increasing but at a diminishing rate, a damping effect on investment needs is produced. There is a need, therefore, for the regular appearance of great new industries, but there is no guarantee that they will appear as needed. At least, it is alleged, none are in sight at the moment.

Finally, when the nation's stock of capital goods ceases to grow rapidly as a result of the factors just described, it is said to become possible for the existing firms and industries to finance their capital requirements to a great extent out of allowances for depreciation, depletion, and obsolescence, as set aside out of earnings. Allowances for these purposes increase in size as the nation's stock of capital increases. The actual machines and equipment for which these allowances are supposed to provide replacements may not wear out for many years, but these allowances are made annually and hence are available for spending for new machines and equipment. Consequently, it is alleged, industries with large allowances for depreciation and obsolescence can modernize and improve their productive equipment, and even at times expand their productive facilities, without having any recourse to the new savings of individuals. Corporations in particular are supposed to be able to finance their capital requirements out of allowances for depreciation and obsolescence, and other earnings, without any need for the "outside" savings of private individuals. Hence there is a continuing problem of finding outlets for these savings of individuals.

A number of special factors operating in the 1920's enabled us to live on borrowed time for a few years. First, there was a great boom in residential construction. We were in the upswing of the building cycle, and the virtual cessation of housebuilding during World War I had created a great shortage which was more than filled in this prosperous period. Second, there was a large volume of public investment in roads, schools, and other public improvements during the period. Third, we had a fairly large outlet for savings in foreign loans and investments which also helped to support our export trade. Fourth, there was an outstanding growth of consumer credit for financing the purchase of durable consumers' goods. Finally, there was a great expansion of the automobile industry and various supplementary industries during the period. After 1929 these supports of our prosperity all collapsed and everything hit us at once. The resulting depression was severe and prolonged, and a really satisfactory level of employment and income was not reached until the heavy deficit spending of the government in the period of World War II raised aggregate demand temporarily to an overwhelmingly high level.

According to the mature-economy theorists, the outlook for the operation of our economic system without extensive governmental intervention and assistance in the postwar period is anything but favorable. They fore-

see total annual savings of close to 30 billion dollars in the United States and cannot imagine what we are going to do with them after the first couple of years of "catching-up" are over. They think that what our economy needs from the government is not an occasional lift when it falls down but a permanent crutch in order that the economy may walk at all. Deficit spending by the federal government, or spending over and above all sums which it takes out of the income of the people, is alleged to be a permanent necessity, and the federal debt will apparently have to go on increasing indefinitely.

CRITICISMS OF THE THEORY OF ECONOMIC MATURITY

Fortunately, from the point of view of those who do not like to see a large degree of governmental intervention in the operation of the economic system, the theory of economic maturity is still only a theory, and it seems to be almost completely unsupported by statistical evidence. All of its main points are open to serious challenge.

The Declining Rate of Population Growth. There can be no doubt that the United States, along with many other countries, experienced an unprecedented growth of population in the late eighteenth and nineteenth centuries, or that the rate of population growth has declined sharply, both here and elsewhere, in recent decades. Moreover, the emergence of a relatively stationary population with an increased proportion of aged people would be certain to have some effects on the composition or pattern of consumer demand, on the mobility of the laboring population between places and occupations, and on the prospective profitability of speculation in urban building sites. However, it is not nearly so clear that a declining rate of population growth will necessarily bring any crisis with respect to the savings-investment problem.

One recent statistical study of many countries over a period of sixty years concludes that "there is no evidence that countries with high rates of population growth have had in general any more rapid rise in their per capita production than others with slow population growth." [5] Similarly, an examination of Germany, Sweden, Great Britain, and the United States during periods of both rapid and slow population growth failed to disclose any tendency for these countries to achieve greater advances in per capita production and standards of living in decades of rapid population growth than in decades of slow population growth.[6] On the basis of these tests, there seems to be no systematic relationship between rate of population growth and the vitality of an economy. Anyhow, the declining rate of population growth is an old friend of seventy years' standing in the United

[5] George Terborgh, *The Bogey of Economic Maturity.* Chicago: Machinery and Allied Products Institute, 1945, p. 40.

[6] *Ibid.*, pp. 40–44.

States. Why should it turn and smite us all at once after living with us in peace for so long?

In the past, economists have commonly thought that, since an economy with a rapidly increasing population has to devote a considerable part of its investment to furnishing the newcomers in the population with their quota of housing, transportation, and productive facilities, it would make slower advances with respect to per capita production and standards of living than an economy with a slowly increasing population which could use almost all of its investment to increase its per capita stock of capital goods. According to the mature-economy theorists, the investment which goes to equip the newcomers in the population with capital goods not only does not compete with the investment which goes to expand the per capita stock of capital goods but is actually necessary in order that the savings of the people may be completely used. Actually, on the basis of the studies mentioned above, the rate of population growth seems to be either neutral or quite insignificant with respect to economic progress.

If this is true, it may be because the reduction of investment outlets in an economy with a slowly growing population tends to bring about a compensating reduction in the volume of saving. In some cases it is undoubtedly true that individuals invest because they save, as, for example, when they place money in a savings account or use it to purchase securities. The situation may be quite different, however, when an individual saves to buy an automobile, a house, or an electric refrigerator, when governments obtain tax revenues to invest in desired public works, when farmers or other unincorporated enterprisers save to buy capital equipment, and when corporations withhold earnings from their stockholders in order to acquire additional capital goods or inventories. In many such cases, if not all, the saving occurs because the investment is contemplated. The point is, of course, that the savings-investment problem may tend to cure itself in an economy with a declining rate of population growth if the volume of savings depends to an important extent on the volume of investment outlets.

One other factor merits consideration here. As a country experiences a declining rate of population growth and aged persons come to make up an increasing proportion of the total population, the volume of dissaving is likely to increase. By dissaving we mean consumption in excess of current income, as financed by the sale or transfer of capital assets to others. Widows expend in consumption the proceeds of insurance procured in the past by the savings of their husbands. Farmers retire, sell their farms, and live in town, spending on current living expenses at least a part of the proceeds from the sales of their properties. Townspeople sell their houses and use the proceeds to pay the rent on less elaborate housing and to make other consumption expenditures. Aged people live on the proceeds of

annuities acquired with savings at some time in the past. Again the point is obvious. To the extent that new savings in an economy with a declining rate of population growth have to be used to permit large numbers of persons to dissave, or make consumption expenditures in excess of current income, any shortage of new investment outlets for savings is compensated for.

The Passing of the Frontier. It is perfectly true that the frontier has disappeared in the United States, but equally true that it disappeared some fifty years ago. Why should its passing wait until 1930 to give us trouble? If the passing of the frontier tended to create a shortage of investment outlets, we should have heard of it long before. As a matter of fact, the ratio of capital formation to national product was slightly higher in the first three decades of this century, after the frontier was no more, than it was in the last three decades of the nineteenth century when the frontier was still with us.[7] And the period 1900–1930 was relatively more prosperous, for the earlier period 1870–1900 contained two protracted and severe depressions. These things cannot be accounted for in terms of a shift from the development of our own frontier to the exploitation of those of other countries, because the United States remained a net importer of capital funds up to World War I. Even after that, we exported capital funds only on a rather moderate scale and only in part to foreign frontier areas.

During the period in which our frontier was being developed, each person who migrated to the frontier merely reduced by one the *increase* in population in an older area. Consequently he would add something to the investment outlets of the country only if he would require more capital on the frontier than it would take to outfit one new person in the settled area from which he came. The actual case seemed to be exactly the opposite, for people in frontier areas were in general very meagerly supplied both with capital goods and with consumers' equipment, and per capita wealth in these areas remained consistently below the level which prevailed in the older areas. The very notion of "frontiering" seemed to involve getting along with very little for a time in order to get in on the ground floor in a new area and prosper with its development. Even if each person on the frontier had required a large complement of capital, no very significant effect on capital outlets would have been observable, for the frontier areas absorbed only about 10 per cent of national population growth in the period from 1870 to 1890.[8] Of course, it can be argued that an important outlet for investment would be created later as the people of the frontier areas not only caught up with those of older areas in wealth and capital but also shared in the general progress of the country in these respects. However, this delightful development, this "afterglow" of the frontier, is an experience

[7] *Ibid.*, p. 65.
[8] *Ibid.*, p. 69.

which we still have ahead of us instead of one on which we should look back with *longing*.

The Dearth of Great New Industries. When the theory of economic maturity was being developed in the late 1930's, it was probably true that no great new innovations or industries were actually in sight. However, even the developers of the theory admitted that this shortage of great industrial developments was not an inevitable accompaniment of economic maturity. Rather it just happened to be present at the time when we were being hit by the effects of economic senility. Even so it was not safe to make predictions as to what might be just around the corner. The future of innovations and new industries is notoriously unpredictable and it is all too common for prophets, when they do not see what is coming, to conclude that nothing is coming.

In any case, it would be easy to exaggerate the importance of great new industries as outlets for investment. It is estimated that the steam railroads accounted for only about 15 per cent of total capital formation in the 1880's, the combined electrical industries (power, telephone, and electric railways) accounted for about 8 per cent from 1900 to 1910, and investment in motor vehicles accounted for about 12 per cent during the 1920's. Even with investment in related industries added in, the automobile industry probably accounted for less than 20 per cent of total capital formation in the 1920's.[9] We always seem to be more dependent for investment opportunities on a mob of small industries than on any great one. And, of course, great new industries provide a net increase in investment outlets only to the extent that they stimulate investment more than they stimulate savings.

It may be conceded that, as the technology of the country develops and becomes more complex, the chance that we shall have innovations as revolutionary as some of the great ones of the past becomes less rosy, while the prospect of large numbers of diversified small innovations becomes better. But this does not mean that technological progress as a whole must be less rapid or less stimulating to investment. Rather, there is hope that this progress will be both more rapid and more steady than in the past. World War II brought scores of innovations, many of which seem destined to become the basis for peacetime industrial developments, great or small. How strange it seems, standing on the threshold of the atomic age, to bewail the shortage of new developments which can furnish an outlet for capital investments!

The Self-financing of Capital Maintenance and Expansion. The argument that, when the nation's total stock of capital goods ceases to grow rapidly, existing firms and industries can finance their capital requirements largely or entirely out of allowances for depreciation and obsolescence, or other earnings, is too complicated to analyze in great detail. However,

[9] *Ibid.*, p. 87.

one investigation concludes that, even if capital consumption (the exhaustion of capital values rather than the actual using up of machines and equipment) "generated fully and automatically the funds for its own replenishment, economic maturity would increase replacement opportunities more than it increased these funds." Again, "the rise of replacement opportunity with economic maturity would compensate in part for the decline of expansion opportunity." [10]

In actual practice, the allowances which are made for depreciation and obsolescence in our economic system as a whole are far too small to provide in full for the capital consumption which occurs. Governmental bodies do not often keep their accounts in a manner which calls for the accumulation of reserves against depreciation, and many unincorporated enterprises (including farm enterprises) are deficient with respect to both general accounting and the making of allowances for depreciation. Individual consumers seldom recognize depreciation on anything except, for income tax purposes, on physical assets held for money income (such as houses which they rent to others). When business firms make allowances for depreciation and obsolescence and the amounts so allocated are not earned in a given period, the allowances merely increase the net losses of the firms from an accounting point of view and do not give the firms funds to invest. Moreover, when the firms do earn the amounts allocated for depreciation and obsolescence, the funds in question are sometimes used for other than capital purposes. Finally, even when the allowances are both earned and used for capital purposes, we have no guarantee that the firms would not have saved similar amounts in the absence of the custom of making these allowances.

The special contention that corporations tend to finance their own capital requirements almost entirely by means of "quasi-automatic" internal savings out of earnings, and thus leave the savings of individuals out in the cold with no apparent investment outlets, is also open to serious objections. In the first place, the argument implies that there is a fixed total volume of investment outlets to be divided between corporate savings and other savings, so that, if more opportunities are used up by the internal savings of corporations, there will be fewer opportunities for the savings of individuals. This is actually far from the case, for investment is a dynamic process and the growth of one firm by means of internal savings out of earnings may lead to the expansion of other firms on the basis of external funds. Thus the Ford Motor Company, once it got under way, developed almost entirely on the basis of internal savings, but indirectly it provided great outlets for personal savings through its stimulation of the expansion of other enterprises and industries.

Again, some corporations are able to finance themselves almost entirely

[10] *Ibid.*, p. 119.

by means of internal savings out of earnings, but others are not. In the period 1923–1929 it is estimated that nonfinancial corporations in this country financed almost three quarters of their capital requirements with internal savings, but in the period 1936–1941 the proportion was only a little over one half.[11] Moreover, during the interval 1931–1934 these corporations depended almost entirely on outside sources for their funds, since their own saving was negligible. On the basis of these estimates there seems to be no trend in the direction of increased financial self-sufficiency on the part of corporations. Corporations have been financing a large (though not necessarily growing) part of their capital requirements by means of internal savings out of earnings for many decades. If this practice has been compatible with a satisfactory solution of the savings-investment problem, and with a high level of national vigor and well-being, for so long, why should it turn malignant all at once?

It is difficult to see why one should say that corporate savings tend to crowd individual savings out of opportunities for investment. If corporations never saved anything but rather paid out all their earnings to the individual stockholders, what now appears as corporate savings would then appear as individual savings (except to the extent that corporations save more for their stockholders than these individuals would save for themselves) and would involve the same total need for investment outlets. Finally, there seems to be no reason to characterize corporate savings as "quasi-automatic," in the sense of being unresponsive to changing conditions and circumstances, for corporate saving declined from around 6 billion dollars in 1929 to less than nothing in 1932.[12] In fact, corporate savings were definitely more flexible than individual savings in this period.

Dangers in the Theory of Economic Maturity. While the theory of economic maturity has a great many followers, both inside and outside Washington, it is necessary to label as both unproven and implausible its case for a persistent shortage of investment outlets in relation to savings in the United States. This does not mean that the theory is not dangerous, however, for if enough people believe in it we shall be impelled to adopt measures which will ensure the appearance of the symptoms of the disease of economic maturity, even though the fundamental ailment itself is not present. If we believe that we are faced with a continuing shortage of investment outlets relative to savings, we are likely to espouse heavy taxes on the higher brackets of personal income in order to mop up funds which would otherwise become "excessive" savings; a combination of high tax and high wage policy to leave firms with only small earnings which may be saved and to increase the consuming power of the masses of wage earners; and continued deficit spending by the federal government to ensure

[11] *Ibid.*, p. 157.
[12] *Ibid.*, p. 156.

the existence of an aggregate demand adequate to take off the market all the commodities and services which our economy can produce at full employment.

There is no doubt that heavy taxes on the higher brackets of personal income will cut into savings, but we should be lucky if these taxes did not affect investment even more severely both by reducing the funds available for investment on the part of receivers of large incomes and by leaving little incentive to invest the funds not taken by the taxes. If our tax policy discourages investment more than savings, it will contribute to the persistent relative shortage of investment outlets which is supposedly symptomatic of the mature economy. The combination of tax and wage policy mentioned above is also dangerous. Individuals have little incentive to operate existing enterprises and even less to found new ones if almost the entire money income of enterprises, over and above the other money expenses of production, must be divided between the workers and the government, leaving little if any earnings for the owners of the firms. In this way we are likely to produce the stagnation which is supposed to characterize the mature economy, even though we do not have the disease itself. Finally, as we shall see in greater detail in the next chapter, continued deficit spending by the government to underwrite aggregate demand and guarantee full employment is likely to have disturbing effects on private business activity and threatens to lead us indirectly if not directly to a planned and controlled economy of fascism or socialism.

QUESTIONS AND PROBLEMS

1. "The proverb which says 'it is an ill wind that blows no one good' applies to changes in the general price level." Do you agree? Explain.

2. "Studying the effects of price-level changes is a waste of time since persons who are harmed when the price level rises tend to benefit when it falls, and vice versa." Show whether you agree.

3. "In periods of depression, falling national income and rising unemployment are opposite sides of the same picture." Explain.

4. "The full effects of depressions cannot be shown by means of statistics." Explain.

5. How do believers in the "business cycle" describe periods of prosperity and depression?

6. "The true explanation of depressions lies in the periodic general overproduction of economic goods." Discuss.

7. "The economy as a whole, in producing a given total value of commodities and services, inevitably generates a national income of equivalent size." Show whether you agree.

8. Explain the self-generating theory of the "business cycle."

9. Do booms and depressions really make up "cycles" in "general business"? Explain.

10. "General *theories* of the business cycle are subject to the same major criticism as general *descriptions* of the business cycle." Show whether you agree.

11. What is aggregate demand, how does it differ from national money income, and what is its significance in connection with prosperity and depression?

12. Explain the significance of underinvestment in connection with the great depression of the 1930's.

13. "Outlays for investment in capital goods are never a large part of total national money income, but apparently they are a very important part." Explain.

14. What is the relationship between a declining rate of population growth and a shortage of investment outlets, according to the theory of economic maturity? Explain.

15. "The passing of the frontier and the dearth of great new industries are important in connection with the modern shortage of investment outlets, according to the theory of economic maturity." Explain.

16. How does the theory of economic maturity account for the great prosperity which existed in the United States in the 1920's? Explain.

17. "A declining rate of population growth does not necessarily cause a persistent shortage of investment outlets relative to savings." Do you agree? Explain.

18. "The theory of economic maturity greatly overemphasizes the passing of the frontier as a factor making for a shortage of investment outlets." Explain.

19. Evaluate the importance of the claim that a dearth of great new industries is causing and will continue to cause a shortage of investment outlets.

20. "When a nation's stock of capital goods ceases to grow rapidly, existing firms and industries can finance their capital requirements out of allowances for depreciation and obsolescence." Show whether you agree.

21. "The special contention that corporations now tend to finance their capital requirements almost entirely by means of quasi-automatic internal savings out of earnings, and thus leave the savings of individuals out in the cold with no apparent investment outlets, is open to serious objections." Explain.

22. "The theory of economic maturity is dangerous because, if we believe that our capitalistic system can no longer operate successfully, we shall be moved to take steps which will ensure that this system will not operate successfully." Explain.

See References for Further Reading at the end of Chapter XXX.

XXX

Economic Instability (*Continued*)

Our discussion of the theory of economic maturity has indicated that this theory is not in any position to give real support to the underinvestment theory proper. We must now return from our digression and consider the strength or weakness of the underinvestment theory itself in relation to the great depression of the 1930's.

THE UNDERINVESTMENT THEORY AND THE GREAT DEPRESSION

Oversaving and Underspending. With reference to booms and depressions, the tendency for investment to exceed saving is shown by an increase in the national income, while the opposite tendency for saving to exceed investment (oversaving) is shown by a decline in the national income. Since we define oversaving in terms of falling national income, it is always present when national income declines, whatever may have touched off the movement. While it is correct to say that a falling national income indicates a tendency to oversaving, or a tendency for investment to fall short of unconsumed income, it is just as correct to say that it indicates a tendency to underspending, or for spending for consumption to fall short of uninvested income. What is true in both cases is that, for some reason, total spending falls short of total income.

Since oversaving or underspending is shown by a declining national income, it is clear that there could have been no oversaving or underspending in 1929 before the decline in national income began. Moreover, it is clear that, once the depression began, oversaving did not *cause* the decline in spending, for it *was* the decline in spending. Oversaving during the depression is the thing which requires explanation. It is not itself the explanation of anything.

Underinvestment or Underconsumption. As we have shown, oversaving or underspending may rest upon either underinvestment or underconsumption, or both. All we really know is that total spending for both consumption and investment tends to fall short of total income in a depression period. Unless we have some way of deciding independently what

the rate of either type of spending should be, we are not in a position to say which factor was responsible for the total shortage of spending. However, the underinvestment theory seems to give credit to a falling off of spending for investment in initiating periods of depression and falling national income.

In the case of the great depression which began in 1929, it may have been true that underspending was initiated largely by a decline in spending for investment, but it was not exclusively so initiated. The stock market crash of 1929 froze many individuals out of their margin accounts and resulted in a net reduction of 4 to 5 billion dollars in brokers' loans in a period of two months. It seems reasonable to suppose that the expenditures of such individuals for consumption may have been sharply affected. At any rate, it has been noted that sales of passenger automobiles in the United States fell off by more than one third in the first three months following the stock market crash of 1929.[1]

Again, it is folly to think that, once the great post-1929 depression got under way, its lengthening and deepening depended solely on a continued tendency of spending for investment to fall off. Every great depression depends upon both underconsumption and underinvestment for the dynamics of its downward spiral, once the movement is initiated. If, in the case of the great depression, we *could* show that spending for consumption had been an entirely passive factor and that a changing rate of spending for investment had been the only active factor, could we then *explain* the depression adequately by referring to underinvestment? Certainly not. All we would have accomplished would be to indicate the general area within which it would be worth while to search for causal factors. We would still need an explanation of what caused the decline in spending for investment.

The Beginning of Depression. One rather popular explanation of the original decline of spending for investment holds that, as production and national income increased in the boom period of the late twenties, expenditures for consumption increased also, but not as much. And this result occurred in spite of the great increase in installment credit for financing the purchase of consumers' durable goods. Because of the unsatisfactory increase in expenditures for consumption, the marginal efficiency of capital (or the rate at which expected future returns from the use of newly produced capital assets have to be discounted in order to make their sum equal to the cost of the assets) eventually had to decline. As soon as this marginal efficiency of capital fell below the rate of interest at which funds could be obtained for investment in new capital assets, the inducement to invest in such assets disappeared.

If there had been a sharp increase in expenditures for consumption at this point, all might yet have been well, but nothing of the sort occurred

[1] George Terborgh, *op. cit.*, p. 185.

because the propensity of people to consume is relatively invariant in short periods of time. With saving tending to outrun investment, the interest rate eventually fell, but it did not fall to zero and hence could not stimulate investment greatly. Moreover, it is held that a falling rate of interest does not greatly encourage expenditures for consumption or discourage savings. Hence oversaving or underspending occurred, a large deflationary gap appeared, and the depression was on. Falling wages and prices did not tend to restore equilibrium because the total national income fell as fast as the total value or cost of commodities and services to be purchased, and even with the smaller national income there was still oversaving or underspending to force the national income down still further.

Eventually capital consumption and other factors will bring such a depression to an end and induce some recovery, but, unless something happens to increase sharply either the inducement to invest or the propensity to consume, balance may be restored at a relatively low level of production and employment. There is no longer oversaving or underspending, and the national income tends neither upward nor downward, but there is chronic unemployment of men and productive resources. This, then, is the famous "underemployment equilibrium" with which it is alleged we would have been afflicted in the 1930's in the absence of governmental measures to stimulate recovery.

The Question of Ultimate Causation. The search for the ultimate causes of anything in the field of economics is usually very discouraging, and this is especially true in connection with economic instability. If expenditures for consumption did not keep pace with production and national income during the boom period of the 1920's, why was this true? Would it be true in other cases as well, or was it unique in the period under discussion? Why should expectations concerning the future returns from new capital goods become so dismal that only a zero rate of interest on borrowed funds could be expected to induce investment? Why should a decline in the rate of interest, or the inducement to save, be almost entirely ineffective in discouraging saving, and why should the combination of falling interest rate and falling prices of commodities and services have almost no effect on the willingness of individuals to make expenditures for consumption?

If we attempted to pursue these questions further, we should probably find partial answers in some peculiar quirks of human psychology, but these, too, would only be partial answers and there would be many things that we would still want to know. For example, was not the inadequacy of the increase in expenditures for consumption in the boom period influenced to some extent by an increased prevalence of monopoly, partial monopoly, and monopolistic competition in the economic system, which made it possible for many enterprises to insist on high prices for their products in the

face of rapidly expanding production and relatively stable wage rates for the great masses of workers? American agriculture during the twenties was burdened with excessive capacity traceable to great expansion during the period of World War I, and farmers shared only to a relatively slight extent in the great prosperity of the boom period. Did this have no effect on the relatively inadequate increase in expenditures for consumption? If the degree of inequality in the distribution of income increased during the boom period, as some people think, was this unrelated to the under-expansion of consumption expenditures?

In view of the complexity of the question of causation in connection with the end of a great prosperous period and the beginning of a severe depression, it is perhaps not surprising that many economists are still unwilling to accept the relatively simple theory we have presented and believe that an adequate explanation must deal with many factors. Among the things to be dealt with would be distortions in the cost-price structure, changes in the relationship of interest rates and profits, maladjustments in particular industries, swings in speculative activities, upward and downward movements of business confidence, variations in the supply of money and credit, changes in inventories, changes in the personal or family distribution of income, modifications of governmental policies, changes in relative debt burdens, and others. Above all we should not overlook the influence of government in prolonging and intensifying the boom period of the twenties.

Twice during the boom period, in 1924 and again in 1927, our economy was apparently ready for a depression of considerable size. In each case, however, the federal government stepped into the breach with measures (to be examined later) calculated to stop the downswing of depression and to induce the resumption of the boom. In the second case, at least, the governmental medicine was apparently too strong. Its use was followed not only by increasing production, employment, and national income but also by an intense, feverish, speculative outburst in the commodities and securities markets. The longer the prosperous period went on, however, the more likely it became that some day the governmental medicine would prove ineffective and that we would be in for a severe depression.

The actual severity of the great depression may have been, at least in part, a matter of chance. If we start two snowballs rolling down from the top of the same hill, it is quite likely that one will go farther and become larger than the other. And so it apparently is with depressions. Almost every great depression misses being a small depression by only a slender margin. In the course of a great depression there will usually be several occasions when the downward movement has almost stopped, but some unforeseen development will come along to give it additional momentum. Such developments, of course, cannot be predicted by any general explanation.

The Incomplete Recovery. It is also necessary to account in some manner for the incompleteness of the recovery which we experienced after the severe depression of 1929–1933. To be sure, we have been told that, unless something drastic happens to increase the inducement to invest or the propensity to consume, we shall have only a partial recovery from depression and may have to get along for many years with a relatively low level of national income, production, and employment. The question then becomes one of why the inducement to invest was apparently so limited in the later years of the 1930's. Since 1929 the construction of housing facilities had not kept pace with the growth of population, and there had been no increase in productive facilities at all commensurate with the growth of the labor force. More innovations had been made and more patents issued in the thirties than in the twenties. Corporations were more dependent upon outside funds than they had been during the boom period. And an unprecedented amount of depreciation and obsolescence of capital goods had accumulated over the years of depression and mild recovery. The physical and technological condition of the economy seemed very favorable for extensive investment, and yet we still had over nine million unemployed in 1939, with a national income well below the level of 1929. In explaining such a situation, it would seem necessary to emphasize the acts and policies of the federal government during the period of the 1930's. This we shall do a bit later.

THE QUEST FOR FULL EMPLOYMENT

Whatever may have been the causes of the great depression or of our incomplete recovery from it, there seems to be a widespread determination on the part of the people of the country that these things must not be allowed to happen again. In view of the innumerable factors which may play a part in bringing on depression, it is not surprising that most proposals for preventing depression and unemployment emphasize not attempts to deal directly with these factors but rather reliance upon the federal government. When it appears that the aggregate demand resulting from private expenditures for consumption and investment plus normal governmental expenditures is going to be inadequate to take off the market all the commodities and services which our economy can produce at full employment, the federal government must step in to underwrite aggregate demand and guarantee full employment. By making public expenditures in excess of the amounts which it subtracts from the incomes of the people, or by other methods, it must keep aggregate demand at such a level that there will be continued full employment for men and other productive resources.

What Is Full Employment? Full employment has developed rapidly into one of those social objectives which everyone must favor, at least in general, or run the risk of being regarded as unsocial, but it is really only

another high-level abstraction. When we trace it down to the level at which we encounter actual flesh-and-blood workers, what does it mean? According to some advocates of a full-employment policy, the term refers to a situation in which everyone who is willing and able to work will also be able to find employment. This hope of having all available workers employed all the time is one which could hardly be realized through governmental policy or in any other way. At any given time, there will always be some workers who are "between jobs," or who have given up certain jobs and are taking some time off before getting busy on another. Some workers are always moving from one place to another in the process of changing occupations. In spite of all that can be done to stabilize production and employment throughout the year, some seasonal variations in production and employment will remain in various industries and businesses as long as we cannot eliminate the seasons themselves. No matter when we survey the economic system as a whole, we shall always find some workers out of employment.

The concept of full employment which we are discussing also seems to assume that it is desirable from a social point of view to have everyone employed who is both able and willing to work. However, during the war period we had mothers working in factories while their young children, when not in school, had to shift for themselves at home or in the streets. We had young people working who should have been in school, and we had old people working who should have been retired under normal conditions. It seems highly optimistic to think that all of these workers will retire voluntarily from the labor market in the postwar period. If they do not, and are still both able and willing to work, should they have employment?

Full employment has also been defined to mean a situation in which there will be more job-vacancies than applicants, although some individuals will be out of work at any given time. Such a situation might be regarded as one of super full employment, except for the presence of some unemployed persons. Again, full employment may mean holding unemployment down to the irreducible minimum, numbering perhaps one to three million persons at a given time. Indeed, some people seem to think that a "float" of five or six million unemployed would not be inconsistent with the operation of a full-employment economy.

The meaning which we assign to the objective of full employment will depend to some extent upon whether we think of it as a conditional or as an unconditional objective. If it is a conditional objective, then presumably we want as much of it as we can get without sacrificing other objectives which we also regard as important, such as individual economic and political freedom, a high rate of economic progress for the economic system, or a satisfactory level of real income. On the other hand, if full employment is an unconditional objective, then we want it even though,

in order to get it, we have to give up our economic and political freedom, get along with an unprogressive economy, or sacrifice some of the real income which we could otherwise obtain.

If full employment is not to be an unconditional objective, we might define the desirable level of employment as the highest level which can be attained without direct governmental control over the economic system. However, such a definition would not be acceptable to many of the wilder advocates of full-employment policy, who seem to think that, if the thermometer indicates that the temperature in the house is too low for comfort, we should not hesitate to apply the governmental torch to the thermometer, regardless of consequences, in an effort to secure a more satisfactory reading. That is, even though unemployment on a large scale is merely a symptom of certain fundamental disorders which exist in the economy, they think we should go directly to work on the matter of unemployment itself.

It is too early as yet to know what the final policy of our federal government will be with respect to the objective of full employment, but the policy embodied in the full-employment bill passed by the Senate in the latter half of 1945 was rather moderate. According to this bill, the President of the United States, in addition to preparing the regular budget of the federal government, would have to turn out annually another document called a "national budget of production and employment." This budget would estimate the total expenditures by consumers, private enterprisers, and government which would be necessary to assure a full-employment level of production in the ensuing fiscal year (necessary aggregate demand), and also the total expenditures by these same persons and agencies which are actually in prospect for the same period (prospective aggregate demand). If prospective aggregate demand is estimated to fall short of necessary aggregate demand, the president is to suggest measures for handling the unemployment problem which is foreseen. However, he is also expected to show where the money is coming from and to provide for a balanced budget over a reasonable period of years, so that the employment-creating expenditures of the government will not add to the public debt.

Disappointed advocates of a more stringent full-employment policy contended that this bill merely guarantees everyone the right to a job if he can find one, and provides that the government may help out if such action is found to be convenient. If, as would seem readily possible, the operation of such a program would permit a considerable volume of chronic unemployment to exist, we may be sure that a strong demand would develop for a full guarantee of employment by the government and for whatever amount of governmental expenditures might be necessary to effectuate the policy.

Do We Need a Governmental Full-Employment Policy? Before discussing the implications of a stringent full-employment policy on the part

of the federal government, we must ask whether it appears that such a policy is needed. Contrary to popular opinion, it is easy to develop a strong negative answer to this question. The record of our capitalistic system in providing a high level of employment, as well as increasing production, national income, and scales of living over the years, is remarkably good. Since 1900, our freely operating capitalistic system has encountered only two depressions in which there was a disturbing volume of unemployment. In 1908, unemployment rose to 2,300,000 or 6.3 per cent of the working force, but in 1909 it dropped back to 719,000. Again in 1921, annual average unemployment mounted to 4,754,000 or 11.2 per cent of the working force, but our freely operating capitalistic system, left once more to its own devices, soon worked things out. By 1922, unemployment had fallen almost in half and by 1923 we were back to full employment with a "float" of only 749,000 unemployed.[2] Such a record over a period of thirty years would be difficult to improve.

But, it may be asked, what about the great depression of the thirties and the incomplete recovery that followed? Did the events of this period not show that our capitalistic system has reached a point where it is no longer capable of operating satisfactorily on its own power? The answer is that neither the boom period of the twenties nor the severe depression and unsatisfactory recovery that followed can be charged against a freely operating capitalistic system. In the first place, the government deliberately prolonged and intensified the boom period. In late 1923 and early 1924, the Federal Reserve System bought some 400 million dollars' worth of government securities in the market. This gave the commercial banks additional reserves, drove interest rates down to lower levels, flooded the money market, and led to an expansion of over 4 billion dollars in commercial bank credit. Another result was the making of foreign loans to European countries which permitted them to import our products in spite of their inability to sell goods to us over the high tariff barrier that we had erected after the war. The strong speculative movement in the securities markets also got under way.

When depression threatened in 1927, the cheap-money policy of the government and the Federal Reserve System was renewed and intensified and the boom period went on, with rapidly expanding bank credit, sharply rising security prices, and large issues of new domestic and foreign securities. We continually increased the obligations of other countries to make payments to us (that is, export goods to us), while our high tariff stood as an insurmountable barrier to such payments. The boom period came to an end with the crash of the stock market in late 1929. If the government had lowered the tariff and otherwise had let things alone, we might have had an ordinary, though fairly severe, depression and revival.

[2] *Financing American Prosperity*, pp. 10–11.

Instead, the government tried to use its same old magic. The Federal Reserve System purchased a large volume of government securities in the market, increased commercial bank reserves, drove interest rates down, and revived speculation in the stock market. This time it did not work. To be sure, the stock market rose in early 1930 and the volume of new securities increased, but the trend of industrial production and employment continued to be downward. Then, late in 1930, Congress enacted a new and higher tariff, which provoked other countries to retaliation, and international trade almost disappeared. And all the time the government was exhorting business leaders to maintain prices and wages at high levels and to increase their outlays for investment in order that purchasing power or aggregate demand might be preserved. No wonder the depression was severe!

Just after the worst of the depression was over, the government decided that recovery was not going to occur "naturally" and that a host of governmental measures would be necessary to induce recovery. Therefore the government increased its manipulation of the money market and the interest rate. It abrogated the gold clause in bonds and other contracts and devalued the dollar in terms of gold. These actions could hardly have increased confidence in our dollar either on the part of American businessmen or on the part of the people of other countries. The government manipulated the prices of agricultural and other products and regimented the activities of agriculture, industry, and trade under the AAA and the NRA. It interfered with wages and the operation of the labor market in general under the NRA, the Wagner Act, and the Fair Labor Standards Act. Under the Securities Act of 1933 and the Securities Exchange Act of 1934, it purified the issuance of securities and the operation of the security exchanges and also greatly restricted activities in these fields.

The government raised taxes in the face of adverse business conditions and used taxes for revising the distribution of income as well as for revenue. It borrowed heavily for financing its expenditures "for the stimulation of recovery" and greatly increased the public debt. The effects of these devices on business confidence, on the desire to operate existing enterprises, and on the willingness to found new enterprises have been commented upon elsewhere. The government also introduced tremendous confusion into the operation of our economy. Governmental policies were subject to sudden and drastic changes, and governmental agencies were created in almost endless succession. In many cases Congress delegated blanket powers to the executive department of government, and executive agencies had or assumed the power to make rulings with the force of law. Reports, forms, and questionnaires, often to be filled out in quadruplicate or quintuplicate, were issued in profusion by governmental agencies, with orders to businessmen to prepare them according to instructions "or else." Businessmen could hardly determine where they stood from one day to another.

Governmental policies were often in conflict with each other during this period. The AAA was to help the farmers by raising the prices of farm products, but the NRA tended to raise the prices of commodities and services which the farmers had to buy. Rising prices under both of these policies tended to interfere with the desired revival of export trade. As already noted, the desire to reform the security markets and the issuance of securities interfered with the desire for a revival of activity in these fields. Under all of these conditions, it is not surprising that recovery from the depression was incomplete in the years from 1933 to 1939. Governmental interference, rather than economic maturity and a dearth of investment outlets, was the source of our difficulties. Under extensive governmental interference with the operation of a capitalistic economy, depressions are likely to be long and severe and recoveries weak and anemic.

According to this point of view, if we wish to have our country achieve a high level of national income, production, and employment in the post-war period, all we have to do is to see to it that the government "gets out of the way." The tax system should be revised not with the idea of reducing governmental revenues but in order to provide incentives for the investment of savings, for the full operation of existing enterprises, and for the foundation of new enterprises. Except in the case of certain "natural monopolies," the government should stop tinkering with prices. That is, the capitalistic price system should be set free so that it can perform its usual functions in a capitalistic system once more and so that prices and costs can be relied upon to tell the truth about underlying market conditions. The public debt of the federal government should be first stabilized and later gradually reduced. The national tariff barrier should be lowered considerably and state interference with interstate commerce should be removed. Most governmental interferences with individual branches of economic activity should be abandoned. In particular, the regulation of the security exchanges and of the issuance of new securities should be toned down. To these suggestions, some people would add a comprehensive and thoroughgoing governmental campaign against monopoly, certain positive measures to stimulate the foundation of new small enterprises, and efforts to develop higher business standards and codes of ethics.

We should not suppose that, under such a program, complete stability of economic activity could be achieved. Depressions would undoubtedly still occur from time to time. In a free economy, reaction, readjustment, and liquidation will sometimes be necessary and will produce temporary unemployment. However, we would be entitled to hope that the depressions which occurred would be short, and that, after each one, we would recover to levels of income, production, and employment higher than those which we had known before. We should remember that the only really unsatis-

factory recovery from depression that we ever had was the one that the government undertook to produce for us in the 1930's.

Full-Employment Policy and Private Enterprise. At the present time it is probable that the majority of the people of this country would not be willing to accept the above analysis of the need, or lack of need, for a governmental policy of underwriting aggregate demand and guaranteeing full employment. The opinion is widespread that we cannot afford to trust the future of our country to the automatic functioning of the capitalistic system, and that we must depend upon governmental direction and control. This opinion is entitled to respect even from those who do not agree with it, for it seems to represent an honest choice between capitalism and the planned economy. On the other hand, some people argue that governmental control through the mechanism of a full-employment policy is necessary to ensure the successful operation of a capitalistic system in this country in the future. This opinion is much more dubious, for it seems likely that the policy in question will result in the destruction of the capitalistic system and will ensure the future existence of a planned and controlled economy.

If the demand for full-fledged underwriting of aggregate demand and guarantee of employment by the government is to be met, will it be possible for the government to make heavy expenditures in the process of creating employment without encroaching on the fields which are traditionally reserved for private enterprise? Some advocates of full-employment policy apparently have no worries on this score. The necessary spending by the government, it is said, would go on in such fields as education, public health facilities, nutrition, slum clearance and housing construction, and resource development (more TVA's), where the need for public outlay is great quite apart from the question of full employment and where there would be little direct competition with private enterprise.

However, the need for expenditures in these fields is not unlimited, and governmental spending for full employment might have to amount to 10 or 15 billions of dollars in some years, or even year after year if, as some people think, there were a chronic shortage of private spending for consumption and investment in the face of this governmental full-employment policy. In such a case it would seem that the government might have to fall back on the multiplication of post offices and courthouses, if not on leaf-raking and hole-digging and refilling, if governmental spending were not to interfere directly with private enterprise. On the other hand, if governmental policy in providing employment did involve direct competition with private enterprise, it might well discourage more employment-creating outlays than it furnished and lead to the gradual replacement of private enterprise by governmental enterprise.

Even if it involved no direct competition with private enterprise by the government, the full-employment policy might still affect private enter-

prise adversely and it might turn out to be a difficult policy for the government to administer successfully. At best the policy would involve a fearful problem of estimation, and this problem would be complicated by the fact that we cannot assume that private spending for consumption and the spending of private enterprises for investment are given entities to which we could simply add governmental spending to provide employment. Suppose, for example, that we are unwilling to increase the public debt any further, that we want the government to derive its funds for employment creation from taxation, and that large total expenditures will be necessary. Heavy taxes, even though levied progressively, will cut into spending as well as so-called unnecessary saving. Individuals in the higher tax brackets may be rendered more unwilling than ever to put further capital funds into risky business ventures, since any gains which they make will be largely taken by the government while any losses which they suffer will be almost entirely their own. Because of the necessity of paying heavy taxes on income, any capital funds actually seeking investment may be channeled almost entirely into fields where the rate of return temporarily promises to be high rather than other fields where it seems likely to be moderate or low. This may produce overinvestment in some industries and businesses and underinvestment in others, with unfavorable effects on employment in the end. In view of such considerations, how could governmental leaders ever decide how much additional tax revenue should be collected and spent in order to provide a given net increase in total employment?

If the governmental guarantee of full employment is to be carried out on the basis of deficit spending rather than taxation, the problem is no less complicated. Heavy deficit spending may have an unfortunate effect on business confidence and arouse fears of higher taxes later on, with adverse effects on the amount of employment furnished by private industries and businesses. Some prices may be more responsive than others to increases in the total volume of spending, and deficit spending may therefore cause changes in the relationships which prevail between the prices of different economic goods. Such distortions of the price structure may then have a reflex influence on the volume of private spending and employment creation. Again, deficit spending, besides providing employment, may raise the general level of prices, and there is no reliable way to determine in advance just how the influence of deficit spending will be divided between furnishing employment and raising prices. However, it is likely that the latter influence will increase and the former will decrease as the economy gets nearer and nearer to full employment.

Not all unemployment is of the "lack of enterprise" or "shortage of aggregate demand" type which we might hope that deficit spending would relieve. Some unemployment is of a type which has been called "structural"

or "market imperfection" unemployment, and results from such things as existing distortions in the price or wage structure, seasonal fluctuations in demand, frictions in the labor market, and abnormally low efficiency, reliability, or adaptability on the part of workers.[3] Deficit spending would be expected to have very little, if any, favorable influence on this type of unemployment. Declines in private spending may be sudden, large, and highly unpredictable, and we may well wonder whether a governmental program of deficit spending would be sufficiently flexible to compensate for such changes in private spending and to maintain a steady level of employment in general.

In the light of such considerations, it would seem doubtful that any human wisdom would ever suffice to determine just how much deficit spending should be undertaken at any given time in order to produce a specified net increase in total employment. Presumably the deficit spending engaged in by the government in the 1930's was intended to induce recovery and achieve a high level of employment, and yet we know that unemployment continued on a large scale throughout the period. It is easy to look backward and decide that this spending program was ineffective because it was carried out on too small a scale, but much more difficult to decide what scale of deficit spending would be adequate to ensure full employment a year or so in advance. High and low points in production and employment in our economy are easy to detect several years after they have occurred. It is quite another problem to determine them as of about the time when they occur, and practically impossible to determine them in advance, as would seem necessary if governmental expenditures for creating employment were to be based on an orderly, planned program. If, as is sometimes suggested, there is no one in Washington who fully understands even the federal government itself, how can we expect any man or group of men to understand the complex economic life of the whole country?

Full-Employment Policy and the Planned Economy. It seems to follow that any estimates of governmental expenditures necessary to provide full employment, whether the program were to be based on taxation or deficit spending, would be likely to be wrong in practice. If the full employment scheme were tried over a period of several years, the estimates of necessary governmental expenditures were always wrong, and chronic unemployment still troubled us, what would be the result? The program might be abandoned as impractical, but we may doubt whether the government would want to give it up and whether it would be allowed to do so. The government might confess its inability to provide full employment in our type of economic system and adopt the objective of maintaining merely as high a level of employment as possible. However, a half-hearted governmental program would not be likely to produce as high a level of employ-

[3] *Financing American Prosperity*, pp. 306–307.

ment as our economy could furnish in the absence of a governmental employment policy.

The government might decide to employ a "shotgun" policy, figuring that extremely large deficits, such as those that existed during World War II, could hardly fail to provide full employment. However, the government's credit in a capitalistic economy and in peacetime could not long sustain deficits of 50 billion dollars per year, and in any case the policy would be likely to produce a runaway inflation as well as, or instead of, full employment. The most likely notion of all is that the president and his advisers would decide that full employment could be maintained only if the government had the power to plan production, employment, wages, prices, and virtually everything else, for the economic system as a whole. Anyone can take it from there! The planned and controlled economy would be at hand.

This result seems to be implicit in some of the arguments advanced by advocates of the full-employment policy. It is said, for example, that our experiences in the war period proved that it is unnecessary for us to have unemployment in the United States. We can surely agree with this proposition. We can readily have full employment in the United States if we are willing to continue to live in a planned and controlled economy such as we had during the war period. Incidentally, many people would consider our wartime economy as intolerable in peacetime even if it worked as well then as it did during the war. And there is a strong presumption that it would not work as well in peacetime. The objectives of an economy in a major war are rather simple and straightforward, and governmental control and coordination may accomplish something toward their attainment. In peacetime, however, the problem would be to produce the thousand and one goods that people want, in the right quantities and at the right times, and to have them in the right places and for sale at prices that will clear the market. There would be no chance then to let civilians get along on what was not needed for war production.

How Would We Like the Planned Economy? All this does not mean that full employment guaranteed by the government is either a "good" objective or a "bad" objective, or that we should or should not pursue it. We have merely been surveying some of the reasons for thinking that, while a half-hearted policy for maintaining employment would probably be ineffective, a strong governmental full-employment policy may be inconsistent with the operation of a capitalistic economy and would be likely to involve us in a planned and controlled economy. Now a planned economy is not necessarily bad. Such a system could certainly operate more or less successfully. It could furnish us with some things that are difficult, if not impossible, to obtain under capitalism, and it would probably deprive us of

some of the things which we are accustomed to having under capitalism. Whether or not we would like it is up to us.

A planned economy could provide a degree of economic stability which we have never known under capitalism. It could keep the wheels turning and the factories producing, and it could avoid those great swings of boom and depression which have marked the operation of our capitalistic system. It could provide full employment according to almost any definition which was considered suitable and, within wide limits, it could devote almost any share of the national income to the purposes of social security. Without furnishing any guarantee as to the size of the total national real income, a planned economy could certainly provide for a more nearly equal division of income among the individual citizens than that which results naturally from the operation of a capitalistic system. Money incomes, of course, could be made as large as desired, for the benefit of people who seem to find a peculiar satisfaction in taking in a large number of pieces of money for their services in a given period. On the whole, if we rated stability, security, and relatively equal division of income rather high on our list of economic preferences, we might like to live in a planned economy. Such a system would probably appeal to individuals who attribute their present lack of success to the evil workings of the capitalistic system rather than to their own inadequacies, who are convinced that the world owes them a living, and who are chronically unwilling to have their welfare depend on competition with other men in the market.

On the other hand, a planned economy could not guarantee everyone a satisfactory real income merely by dividing the total income of the economy rather evenly among the citizens, while keeping the factories and people at work. It is necessary to have a reasonably large national income to divide before the shares, however nearly equal they may be, will also be adequate. The size of the total real income to be divided in a planned economy would depend upon the incentives of the people, other things being equal, and such an economy might have a serious problem of incentives. The problem would really have three phases. First, how do you get people to work? Second, how do you get people to work hard instead of merely going through the motions? Third, how do you get people to undertake the most important, difficult, and responsible work for which they are qualified? A planned economy, with compulsion as well as positive incentives at its disposal, could probably solve the first phase of the problem, but it might have more trouble with the second and third phases in view of its guaranteed employment, full social security, and relatively equal distribution of income.

Again, we should be quite certain to experience a loss of freedom in the planned economy, and this matter also has several phases. The planned economy can furnish freedom *from* things—freedom from want and free-

dom from fear (except fear of the government)—but very little freedom *as such*. The individual would no longer be free to found his own enterprise in any field that pleased him and to produce and offer for sale as much or as little product as he wished. He would no longer be free to own land and capital and to devote them to whatever purpose seemed likely to be most productive of income. He would not be free to seek to maximize his income in competition with other men and to hold fast to what he could acquire.

The individual as a worker might also lose freedom, for too much security may not be good for labor efficiency. If jobs at good wages must be made available to all, what will keep the worker from taking a week off to go fishing when he pleases? If he feels below par in the morning, what will keep him from showing up for work at nine or ten o'clock instead of at eight? What will keep him from changing his job very frequently in a search for greener pastures? Serious problems of these kinds have arisen in the Russian planned economy of full employment and the result has been severe limitations and penalties relating to absenteeism, lateness to work, and labor turnover, and even a labor draft to get people into various necessary but unpleasant lines of work. If similar devices were adopted in our planned economy, the worker would find that he had reached security of employment at good wages only by being tied hand and foot in his job. He would learn that, if the government assumes responsibility for a person's welfare, it is very likely also to assume responsibility for that person's conduct. And he might come to suspect that employment is only a means to an end rather than an end in itself.

It is likely that labor organizations would lose significance in the planned economy. If full employment is to be guaranteed at high wages and under good working conditions, the government here, as in Russia, probably would not relish attempts on the part of labor unions to secure still higher wages and better conditions of employment by collective bargaining. As a result, the unions might well degenerate into organizations for carrying on social and recreational activities, for handling grievances within enterprises, and possibly for the detailed administration of social security funds. While the government might require the workers to be members of unions, these organizations and their leaders would be much less important than in our present economy.

Finally, if we were to have a planned and controlled economy, there would be another aspect of freedom to worry about. In theory, if the economic planners did not give the people what they wanted in a given period of (say) four or five years, the citizens would rise up at the next election, throw the rascals out of office, and substitute another set of planners who would make plans which were better suited to the desires of the people. But this might be easier said than done, for there is a grave

question as to whether full-fledged economic planning is compatible with the democratic process. Economic planning requires that enormous powers be concentrated in the hands of certain governmental agencies which are charged with the making and enforcement of the economic plans, and these great powers may be used to destroy democratic control of the government. The planners may come to think that they know what the people need and desire better than the people themselves know, and may be entirely unwilling to doff the great powers with which they have been clothed. It is a common experience to find that it is far easier to give great powers to governmental officials than it is to get them back. In such a case one of the principal plans on which the economic planners would work would be a scheme for keeping themselves in office. In practice, the planned and controlled economy is more likely to be associated with dictatorship than with democracy in government.

If most of the people of this country really believe that the advantages of the planned economy outweigh the disadvantages and want to live in that kind of system, they should have the opportunity to do so. But the question should be decided in the open. We should not allow ourselves to be convinced (1) that the government can assume full responsibility for individual welfare without interfering in any way with the individual's freedom of conduct; (2) that millions of people can be given more income than they can earn while all others receive all they can earn and keep their incentives unimpaired; (3) that all-powerful governmental monopoly is a safe refuge from the ravages of private monopoly; (4) that we can have less governmental control in the future by adopting complete governmental control at once; (5) that we should adopt economic planning at once because our capitalistic economy may operate so poorly that we shall be likely to drift into a planned economy at some time in the future; (6) that we can insure the continued operation of a capitalistic system by setting up conditions under which such a system cannot operate; or (7) that we can preserve the capitalistic system by substituting a planned economy for it. We should not proceed to the planned economy by subterfuge.

QUESTIONS AND PROBLEMS

1. "A declining national income in depression can be interpreted in terms of either oversaving or underspending." Explain.

2. "Underspending in a depression may result from either underinvestment or underconsumption, or both." Discuss.

3. Criticize the notion that underinvestment was the only factor responsible for beginning and continuing the great depression of the 1930's.

4. Present the most popular theory concerning the causes of the decline of spending for investment at the beginning of the great depression of the 1930's.

5. What is meant by an "underemployment equilibrium"?

6. "It is very difficult to determine the ultimate causes of a great depression." Explain.

7. "Many economists believe that any great depression must be explained in terms of a large number of causes." Explain.

8. "The incomplete recovery of the late 1930's can be satisfactorily explained in terms of unfavorable physical and technological conditions." Show whether you agree.

9. "The meaning of full employment is not easy to determine." Explain.

10. Should full employment be a conditional or an unconditional objective of our economic system? Explain.

11. "Our capitalistic economic system has clearly demonstrated that it cannot be relied upon to furnish a high level of employment for its people." Do you agree? Explain.

12. "Neither the boom period of the 1920's nor the depression and incomplete recovery of the 1930's can be charged against a freely operating capitalistic system." Explain.

13. "If we want our economy to achieve a high level of income, production, and employment in the postwar period, all we have to do is to see to it that the government gets out of the way." Show whether you agree.

14. "Large governmental expenditures for the maintenance of full employment would not necessarily result in direct governmental competition with private enterprise." Discuss.

15. "Any governmental estimates of the governmental expenditures necessary to provide full employment, whether the program were to be based on taxation or deficit spending, would be likely to be wrong in practice." Do you agree? Explain.

16. What are the obstacles to estimating the amount of deficit spending necessary to provide a given net increase in employment at a given time? Explain.

17. "If the full-employment policy proved ineffective in our existing system, the result would probably be a movement toward the fully planned and controlled economy." Show whether you agree.

18. What are the economic advantages which a fully planned economy could provide? Explain.

19. "A planned economy would deprive us of some features of a capitalistic economy which we should like to retain." Explain.

REFERENCES FOR FURTHER READING

Estey, J. A., *Business Cycles*. New York: Prentice-Hall, Inc., 1941, Chaps. 4–15.

Financing American Prosperity. New York: Twentieth Century Fund, 1945, Chaps. 1–8.

Haberler, G. von, *Prosperity and Depression*. Geneva: League of Nations, 1940, Chaps. 1–12.

Hansen, A. H., *Business Cycle Theory*. Boston: Ginn and Company, 1932, Chaps. 1, 2, 4, and 6–8.

Mitchell, W. C., *Business Cycles: The Problem and Its Setting*. New York: National Bureau of Economic Research, 1927, Chaps. 1, 2.

Robbins, Lionel, *The Great Depression*. New York: The Macmillan Company, 1935, Chaps. 1–4, 7, 8.

Schumpeter, J. A., *Business Cycles*. New York: McGraw-Hill Book Company, Inc., 1939, Chaps. 1–4, 8–15.

Terborgh, George, *The Bogey of Economic Maturity*. Chicago: Machinery and Allied Products Institute, 1945, Chaps. 1–13.

XXXI

Economics in War and in Peace

In the last few years most economists have probably been questioned many times concerning the effects of war, or the return from war to peace, on the principles of economics. Individuals who are untrained in economics or who are only slightly acquainted with the subject have commonly assumed that the operation of economic principles must have been almost entirely suspended during World War II, and that the conditions which would follow the end of that conflict would require extensive revision of these principles. It is also commonly assumed that the economic problems which we encounter in wartime are not at all like those which exist at other times, but rather are altogether strange and new. The purpose of the present chapter is to investigate and evaluate these attitudes.

ECONOMIC PRINCIPLES IN WARTIME

The Basis of Economic Activity. In the first place, economic activity rests upon the same general foundations in wartime as in peacetime. Economic activity always arises out of the abundance of human wants and the scarcity of the direct and indirect means for their satisfaction. Human wants are always great in relation to available quantities of commodities and services and in relation to available quantities of the productive agents which can produce the desired commodities and services. The fundamental scarcity of the productive agents is even more obvious in wartime than in peacetime. That is, in peacetime many unsatisfied human desires fail to register in the market because they are not backed up by the ability to pay prices for economic goods, and people sometimes gain a false impression as to the adequacy of available quantities of productive agents by relating them only to the effective demands of consumers in the market. In wartime, when the great demands of the government for goods necessary to the prosecution of the war, backed by adequate purchasing power, are piled on top of the effective demands of the citizens for goods for civilian consumption, it is easy to see that available quantities of the pro-

ductive agents are inadequate to produce all the commodities and services which are desired.

The Principle of Opportunity Costs. In wartime as in peacetime the individual agents of production have alternative uses, and the real cost of securing increased quantities of a given good is found in the amounts of other economic goods which must be given up if increased quantities of productive agents are to be devoted to the production of the given good. However, while it is always true that the use of productive agents for a given purpose involves the giving up of other opportunities for their use, the truth of this proposition is brought home to us most forcibly in wartime. In peacetime, when considerable quantities of certain productive agents may be standing idle, it sometimes seems that nothing of importance is given up when quantities of productive agents are devoted to a specific purpose. In wartime, when the production of adequate numbers of airplanes, tanks, guns, jeeps, and bulldozers clearly involves our going without the quantities of automobiles, washing machines, electric refrigerators, and other commodities which we also desire, both the alternative uses of the productive agents and the significance of the principle of opportunity or alternative costs are readily seen.

The Law of Diminishing Productivity. Since they involve no assumption concerning institutional conditions in the economy or the presence or absence of governmental controls, economic principles in the field of production are as valid under wartime conditions as at any other time. Consider, for example, the Law of Diminishing Productivity. The various agents of production are imperfect substitutes for each other, and this means that, if methods of production remain unchanged and equal successive units of some one agent of production are used in conjunction with fixed amounts of other productive agents, the average product per unit of the variable agent, after a certain point has been reached, will decline. This proposition is valid whether the productive agents are privately owned or publicly owned, whether the control of production rests in the hands of private individuals or governmental agencies, whether those in control of production are motivated by the desire for profits or by other considerations, and whether the organization of production is competitive or monopolistic. There is, therefore, nothing about wartime conditions which can affect the operation of this law.

Other Principles of Production. Much the same thing is true of other principles of production. The individual agents of production, besides being imperfect substitutes for each other, are not perfectly divisible. This means that, in any industry, productive enterprises which are large enough to make effective use of the various indivisible productive agents will have greater efficiency and lower costs than smaller enterprises. Since the indivisibility of the productive agents continues to exist in wartime, the validity

of this proposition is unaffected by war. Since it may be easier for the government, in acquiring the goods necessary for war, to deal with a few large-scale firms than with many small ones, large-scale concerns may be artificially stimulated at the expense of smaller firms in wartime, but this again has nothing to do with the general proposition that firms of optimum size are more efficient and have lower costs than smaller firms. Again, specialized production will be more efficient and less costly than nonspecialized production in wartime as in peacetime, and indirect production, which makes use of large quantities of capital, will have similar advantages over direct production.

Saving and Capital Formation. In wartime as in peacetime, large quantities of capital will not be available for use in production unless the process of saving and capital formation functions efficiently. In a money sense, this process involves the spending of part of the money income of the economy, directly or indirectly, for capital goods rather than for consumers' goods. In the nonmonetary sense, saving and capital formation require the allocation of a part of the productive resources of the nation to the production of capital goods rather than consumers' goods. The cost of obtaining capital goods, which will make for increased production later on, is the same at all times. It is found in the necessity of doing without, in the present, the quantities of consumers' goods and services which could have been produced by the agents of production actually allocated to the production of capital goods.

In wartime the government needs large quantities of various kinds of goods for the prosecution of the war, and productive facilities in many industries may need to be extended. Moreover, the large money incomes in the hands of the people and available to be spent for limited current supplies of civilian goods need to be reduced if severe inflationary pressure on prices is to be avoided. There is thus a considerable need for saving (abstaining from consumption) in wartime, even though only a relatively small part of the funds which are not spent for ordinary consumers' goods and services may actually go into capital formation. The government in wartime is unable to allow the individual citizens to decide in what proportions they will save and spend their money incomes, and it reduces these incomes by heavy taxation and the sale of large quantities of government bonds. However, the fundamental meaning of saving and the nature of the cost involved in the process are not affected by wartime conditions.

Commercial and Investment Credit. Modern methods of production require the use of commercial and investment credit in wartime as in peacetime. Since production continues to be specialized, large scale, and roundabout, it consumes a considerable amount of time, and it is usually impossible to make the various expenses which must be met coincide in time with the receipt of payments for the final products. During the productive

process, therefore, someone must "wait"; that is, extend commercial credit by giving up money or goods at present in return for an equivalent value later. In wartime the government may extend short-term funds directly to business firms which are supplying it with war goods, and numerous regulations may be applied to ordinary uses of commercial credit between business firms and individuals, but the process as a whole goes merrily on.

A similar condition exists in connection with investment credit. Whether the prevailing conditions are those of war or peace, the production of such durable capital items as machines and factories involves certain expenses which must be met ordinarily within a comparatively short period of time, but the durable capital goods themselves give off their benefits over a long period of time, and sometimes many years, instead of all at once. Again it is clear that the costs and resulting receipts of income cannot be made to coincide in time and that someone must make advances at the beginning of the process and wait for a long time for reimbursement. In wartime it is quite likely that business firms will receive directly from the government those advances of funds necessary to finance the setting up of new productive facilities, instead of reinvesting earnings or borrowing from individuals and institutions through the sale of bonds, but this does not affect the fundamental nature of the investment credit process.

Principles of Value and Income Distribution. Some at least of the principles which we have studied in connection with the theory of value may be expected to be just as valid in wartime as in peacetime. One example is the Law of Diminishing Utility, which holds that the intensity of a person's desire for a unit of a given economic good varies inversely with the amount of the good which he already possesses, and diminishes progressively as additional units are acquired. This principle is clearly just the psychological counterpart of the Law of Diminishing Productivity, and it rests upon no assumptions as to institutional factors in the economy or the presence or absence of governmental controls. In war or in peace, with a small amount of a good on hand, the strength of the individual's desire for an added unit of the good is relatively great, while with a large amount of the good already in his possession the desirability of an added unit is relatively small.

The Law of Demand in economics states that the quantity of an economic good which buyers are willing to purchase varies inversely with the price, in a given market at a given time. Whenever individuals have money incomes which they may spend on the various economic goods available in the market, the existence of an unusually low price for a good on a given day tends to produce a large volume of purchases, while an unusually high price tends to curtail the number of purchases made. In wartime as in peacetime, bargain days and bargain prices tend to produce large crowds of buyers and an increased volume of transactions. Of course, as prices rise

over a wartime period, individuals may become willing to purchase increased rather than decreased quantities of goods because they fear the necessity of having to pay still higher prices if they postpone purchases until later on, but this has nothing to do with the validity of the Law of Demand as of a given point in time. In general, then, the Law of Demand is valid in wartime. Moreover, the demands for some economic goods would still be predominantly elastic, while those for other goods would be predominantly inelastic.

Many of the cost tendencies which affect the producers of economic goods operate as usual in wartime. Firms which have fixed quantities of certain productive facilities which may be operated at different rates by using varying quantities of other productive factors will have fixed and variable costs as usual. Moreover, both these costs and the average and marginal costs of production will behave in familiar fashion as the outputs of the firms are varied, and will have the usual relationships with each other, although the absolute levels of various costs may be affected in wartime. Finally, the individual firms in any industry in wartime, if they are left free to decide for themselves in such matters, will be willing to expand output with their given productive facilities only so long as the marginal revenue derived from the sale of additional goods is adequate to cover marginal cost of production.

The economic principles of value and income distribution which lose practical significance to some extent in wartime are those which depend for their validity upon the untrammeled operation of capitalistic institutions and the existence of competitive conditions in the markets for economic goods. Consider, for example, the Law of Long-Run Competitive Price, which holds that the price of an economic good in the long run under competition, as determined by demand and supply, tends to equal cost of production per unit for all the firms in the industry. Under private property, freedom of enterprise, economic motivation, competition, and the absence of governmental controls, the existence of a price for a good well above the level of minimum long-run average cost for the firms already in an industry would be expected to lead to the entrance of new firms and productive agents into the industry, an increase in the productive capacity and output of the industry, and a fall in the price of the product to the level of cost of production.

Under wartime conditions, the price of an economic good may rest at a profitable level over a considerable period of time and yet no increase whatever in the size, productive capacity, and output of the industry may occur. That is, the government may be unwilling to have the industry expand, and its restrictions on the use of materials, manpower, and equipment may make it impossible for new firms to acquire the productive agents necessary for entrance into the industry. Indeed, although the prices

of certain products, such as alcoholic beverages or bicycles, may be such as would allow the firms in the industries to operate at full capacity and make plenty of money, in wartime the government may require the firms to curtail production severely, or even stop it altogether from time to time, in order that their productive facilities may be used to turn out products which are more essential for the prosecution of the war. In other industries, firms may be relieved entirely of their usual worries concerning the market price of their product, for the government may agree to buy given quantities of product, or even all they can produce, at a price equal to cost plus a certain percentage of return for the firms. Thus certain industries may operate at full capacity or even undergo expansion in wartime, even though no such developments would have occurred under ordinary conditions.

In roughly similar fashion any economic principle which indicates how the money income of society will be distributed among the owners of the various productive agents under competitive conditions, and in the absence of governmental interference, is likely to lose significance to some extent in a wartime situation in which the government fixes the prices of the products of these agents and also directly controls wage rates, interest rates, and rents. The trouble in all such cases is not that the economic principles are invalid under the conditions which they assume, but rather that these assumed conditions may not prevail in practice in wartime. Firms and industries in wartime will react to favorable or unfavorable prices of their products in the usual fashion to the extent to which they are permitted to do so, but these reactions are likely to be inhibited by governmental controls in some cases. However, as long as productive wealth is privately owned and industries and businesses are privately operated, price-cost relationships cannot lose significance completely even in wartime. Firms and industries are unlikely to go on operating for any great length of time at prices which do not cover costs.

The Law of Comparative Advantage. In wartime as in peacetime, differences in endowments of productive agents exist between countries and lead to differences in the efficiency with which various types of economic goods can be produced from one country to another and hence to variations in the cost of producing these goods. It also remains true that the various countries of the world could maximize production and standards of living by specializing in their more efficient lines of production and engaging extensively in international trade. However, in wartime a country is likely to be cut off from its normal foreign sources of supply of certain economic goods and be forced to rely on other high-cost foreign sources or to carry on production at home as best it can. Moreover, the international trade of a country can be used as a powerful weapon in wartime, and the government is certain to restrict and control it severely for

this purpose. Under these conditions the international trade of a country in wartime may bear little resemblance to the trade which it would have in peacetime and in the absence of governmental restrictions. Thus we should understand that the Law of Comparative Advantage, which states that a country tends to export those economic goods in the production of which it has a comparative advantage and to import those economic goods in the production of which it has a comparative disadvantage, has full significance only when countries are free to trade with each other and in the absence of governmental restrictions on trade.

The Quantity Theory of Money. The Quantity Theory of Money operates very successfully in wartime if the government permits it to do so. The total expenditures which the government wishes to make annually in wartime frequently exceed by a large amount the total revenues which it is able or willing to extract from the current money incomes of the citizens by means of taxation and the sale of bonds. Whether the government finances the balance of its expenditures by printing paper money for the purpose or by turning bonds over to the banks which pay for them by creating demand deposits for the government to spend, the total quantity of money and credit in circulation expands in relation to the total volume of commodities and services available to be purchased. The result is a tendency for prices to rise in wartime, and the inflation of prices, if unimpeded, may reach serious proportions.

However, the operation of this process requires that the people who come into the possession of the increased quantities of money and credit should be free to use them to bid up the prices of things in general, and there is an excellent chance that the government will not permit this condition to exist in wartime. That is, the government may step in to control or fix the prices of most or all kinds of economic goods. In such a case, except for what can be spent for available consumers' goods and services at controlled prices, the increased amounts of money and credit in the hands of the citizens may have to lie idle in their pockets or banks if they do not wish to invest their funds in government bonds. Even here, of course, one could argue that money and bank deposits which cannot be spent for things in general are not really in circulation or that the velocity of circulation of money and credit is sure to be reduced by the presence of large quantities of idle funds, so that the conditions specified by the Quantity Theory of Money are not present in practice. In any case, however, it is certain that increases in the quantity of money and credit in relation to the volume of trade can operate automatically to produce rising prices only if the government does not interfere with the process.

Summary. Our general conclusion is that many of the principles of economics are as valid in wartime as in peacetime, since they do not depend in any way on institutional factors or the presence or absence of govern-

mental economic controls. There are some economic principles which, while valid enough under the conditions which they assume, may lose significance in wartime because the conditions under which they can operate are not present in practice; but on the whole the reports concerning the destructive impact of war on the principles of economics have been greatly exaggerated.

ECONOMIC PROBLEMS IN WARTIME

War Production. In turning to the discussion of economic problems in wartime, we note again that the carrying on of a major war requires the production of tremendous quantities of economic goods. According to an estimate of the United States Department of Commerce, the national product or the total value of goods and services produced in the United States amounted to $197,000,000,000 in 1944. This figure may be compared with one of only $88,600,000,000 in 1939.[1] Out of the total national product in 1944, some $85,000,000,000, or almost as much as the entire national product in 1939, consisted of economic goods purchased by the government for war purposes. The war expenditures of the government presumably included payments to members of the armed forces and their dependents and payments to certain governmental employees, but a large part of these expenditures were for commodities. The commodities needed for war include many varieties of goods. Some things, such as food and clothing for members of the armed forces, trucks, and ammunition for small firearms, are the same as or similar in type to commodities which are produced in peacetime, though the total requirements for such commodities in wartime may be greater than at other times. Other commodities, such as tanks, jeeps, and landing craft for delivering men and supplies on beaches, are scarcely produced at all in peacetime.

How do we go about producing the enormous quantities of commodities which are needed for war? For one thing, we attempt to make full use of our available productive resources and facilities. When World War II began in September, 1939, facilities for producing steel in the United States were being operated at less than 73 per cent of capacity, but by June, 1942, steel production had been increased to over 99 per cent of capacity. Again, more than five million additional workers were put to work in the two-year period from March, 1940, to March, 1942, and the volume of unemployment was reduced to a minimum. When a war comes along at a time when large quantities of productive resources are standing idle, as did World War II in the case of the United States, great increases in production can be brought about merely by setting available productive facilities to work.

In the second place, the production of the commodities needed for war

[1] The actual increase in "physical" national product from 1939 to 1944 was not as great as these figures, expressed in dollars, would indicate, for the prices of things in general increased to some extent over this period.

is increased by converting productive facilities and labor from the production of civilian goods to war production. Early in 1942, for example, the production of automobiles in the United States was brought to a standstill, and our automobile factories were converted to the production of airplanes and airplane engines, trucks, tanks, jeeps, self-propelled guns, and a host of other things. In other industries, conversion to war production was less complete and a certain amount of production of civilian goods was allowed to continue, but a large part of existing productive facilities was changed over to war production. The same sort of development occurred in the case of labor. From late 1941 to the end of 1943, there was an increase of 7,700,000 in the number of individuals in the armed forces, an increase of 16,000,000 in the number of workers engaged in war production, and a decrease of over 12,000,000 in the number of workers in industries producing civilian goods.

Finally, the production of the commodities needed for war is increased by the construction and operation of new productive facilities for this purpose. In about 2½ years after the outbreak of World War II, Congress appropriated 14.4 billion dollars for constructing industrial plants required in war production, and private business firms had devoted some 2.3 billion dollars to the same purpose. The size of this program of plant expansion may be understood by comparing the 16.7 billion dollars allocated to it with the net book value of 22.5 billion dollars which all manufacturing plants in the United States had in 1940. The expansion of productive facilities in wartime may also be furthered to some extent by using for this purpose some funds which would normally be used for the replacement and maintenance of existing productive facilities in civilian industries.

The Conversion and Extension of Productive Facilities in Peacetime. However, we should not think that the conversion and extension of productive facilities are processes which go on only in wartime. In peacetime, too, it is frequently necessary for some industries to expand their productive facilities, while other industries contract theirs, and for productive facilities to be converted from one field of production to another. In a capitalistic economy such extension and conversion of productive facilities occur on the basis of price relationships, as we noted in Chapter IV. If the output of some economic good under competition is small relative to the effective demand for it in the market, the price of the good is likely to rise well above the level of cost of production. Such a favorable price induces the firms in the industry to make full use of existing productive facilities and, if it continues, will lead to the entry of new firms into the industry. These new firms, because of the favorable price of the product, will be able to bid effectively for the agents of production necessary to their operation. Moreover, part of the productive agents thus attracted may come from other industries whose products are subject to declining demand. As productive

facilities wear out and funds are accumulated which would ordinarily be used for their replacement, these funds may be used instead for investment in productive facilities in the first industry.

That is, if the output of certain economic goods under competition is large relative to the effective demand for these goods in the market, their prices are likely to fall well below the level of cost of production. These unfavorable prices induce the firms in the industries to operate at less than full capacity and, if they continue, will lead some firms to leave the industry and transfer productive agents to other fields of production. Thus existing productive facilities are fully used or used only in part, productive facilities are extended in some industries and contracted in others, and facilities are transferred from some fields of production to others, on the basis of price relationships in peacetime. And the process applies to a great extent to workers as well as to other productive agents. That is, new workers are directed into the fields of employment where they are most needed and old workers are induced to transfer from one industry to another on the basis of favorable wages and other conditions of employment.

The Control of Production in Wartime. Theoretically the processes which we have been describing could be allowed to depend on price relationships even in wartime. For example, the government might undertake to pay such high and profitable prices for airplanes that existing airplane factories would run at full capacity, automobile producers would voluntarily curtail or stop the production of automobiles and convert their facilities to the production of airplanes, and new firms would be induced to enter the field of airplane production. Similarly, the industries engaged in war production could attempt to attract the necessary workers by paying wages much higher than the workers could obtain in civilian industries, and so on.

However, in practice the government is likely to resort to direct control over production and over the extension and conversion of productive facilities in wartime. The reason is not that the problems in question are unusual in *nature,* but rather that they are unusual in *scope* and *urgency.* In peacetime the conversion and extension of productive facilities occur gradually and only a small number of industries may be seriously affected at any one time. In wartime, the conversion and extension of productive facilities are needed "right now" and most, if not all, industries are affected to some extent. Moreover, if the government is creating a considerable part of the funds necessary to finance its war expenditures by borrowing from the banks, reliance on price relationships to induce the extension and conversion of productive facilities may be ineffective and costly. If the government spends much more for war than it takes out of the current incomes of the citizens through taxes and direct sales of bonds, the willingness and ability of the government to pay high prices for airplanes may be offset to a considerable extent by the willingness and ability of private citizens to pay

high prices for automobiles. In such a case, the extension and conversion of productive facilities may not go forward as desired, and the money cost of the war may be greatly increased as prices rise because of the competitive bidding of government and citizens for war goods and civilian goods, respectively. When the agents of production are fully employed, nothing further except inflation can be brought about by competitive bidding for their services.

Soon after the United States entered World War II, the task of speeding and guiding the output of necessary goods was assigned to the War Production Board, which was created by two executive orders issued by President Roosevelt. The War Production Board was authorized (1) to exercise general direction over the war procurement and production program, and (2) to determine the policies, plans, procedures, and methods of the several federal departments, establishments, and agencies in respect to war procurement and production, including purchasing, specifications, and construction, and including conversion, requisitioning, plant expansion, and the financing thereof; and to issue such directions in respect thereto as might be deemed necessary or appropriate. The board established a number of regional offices, located in large cities from coast to coast, to facilitate the handling of local and regional problems. The powers of the board were, of course, very great. It could order industries producing civilian goods to curtail production or even stop operating altogether, and it could permit them to resume operation when needs for war production moderated to some extent. It could take almost any necessary steps to insure the adequate production of goods necessary for war. Conversions and extensions of plant facilities could be undertaken only with the approval of the board, and so on.

The allocation of vital materials among firms and industries, under the auspices of the War Production Board, occurred on the basis of the so-called Controlled Materials Plan. Prime contractors all over the country were required to assemble bills of materials needed, specifying kinds, quantities, and the times at which they were needed. These bills of materials were submitted to seven claimant agencies (Army, Navy, Maritime Commission, Aircraft Scheduling Unit, Board of Economic Warfare, Lend-Lease Administration, and Office of Civilian Supply). The claimant agencies combined the bills of materials sent in by their contractors and submitted them to the War Production Board, which matched total requirements of the claimant agencies with available supplies of materials, making allowances to the various claimant agencies according to their importance. The claimant agencies followed their own preferences in distributing their allotments to the prime contractors, who in turn rationed the materials to subcontractors. The allotments to contractors amounted to "certified checks" for specific quantities of materials in specific periods of time. Finally, the War Pro-

duction Board had to inform the producers of the materials as to the kinds and amounts to produce in order to meet the allocated demand in each period.

The handling of labor or manpower problems is also very important in connection with production for war. The manpower problems include (1) the activation of the labor supply (getting women, the aged and the young, and individuals who would ordinarily be in the leisure class out to work); (2) the allocation of manpower among the armed forces, industries engaged in war production, and civilian industries and businesses; (3) the adjustment of hours of work and other matters so as to maximize production; (4) problems of motivation (wages and related matters); and (5) the handling of labor disputes. During World War II, the first three of these classes of problems were turned over to the War Manpower Commission in the United States. This commission was created by President Roosevelt in April, 1942.

Under the influence of high wages, appeals to patriotism, and other incentives, the total labor force of the United States (including members of the armed forces) increased from 54 million individuals in 1940 to almost 64 millions by the end of 1943. The allocation of this labor force among the armed forces, industries engaged in war production, and civilian industries and businesses depended on several devices. During part of the war period, the War Manpower Commission had charge of the Selective Service System and could decide as to the disposal of individuals between service in the armed forces and productive activities at home. At other times, the Selective Service System was independent of the War Manpower Commission, and the commission could make decisions relative only to those individuals who were not needed for the armed forces. In achieving the desired distribution of workers among industries and businesses, workers were "frozen" in essential jobs and were not permitted to leave unless they could get certificates of availability from their employers, workers were notified to get out of a large number of nonessential occupations or lose their deferred status with respect to Selective Service, workers leaving essential occupations were made available for immediate induction into the armed forces, wage differentials between firms and industries, which led to undesirable migration and labor turnover, were reduced, and so on. In increasing the productivity of the labor force (disregarding questions of wages), the length of the maximum work week was increased considerably in war industries and elsewhere, workers were given additional training, and efforts were made to improve conditions with respect to lighting, ventilation, sanitation, air conditioning, sound proofing, factory restaurants, and housing.

In general, we should reach the same conclusion regarding manpower problems in production as in connection with production problems in gen-

eral. Problems of getting a total labor force of appropriate size, of allocating the labor force among industries and businesses, and of maximizing the productivity of labor are always with us in peace or in war. However, while these problems are ordinarily handled to a great extent on the basis of price relationships, their unusual scope and urgency in wartime may lead us to rely largely on governmental intervention for their solution.

Industrial Conflict in Wartime. The objectives of workers and employers are approximately the same in wartime as in peacetime. Workers are interested in getting high wages, short hours, and satisfactory working conditions, while employers are anxious to derive substantial earnings from their businesses. These interests of workers and employers are likely to be in conflict to some extent, and either party may seek to gain advantages which will not be willingly granted by the other party. In war or in peace, the result is likely to be the use of strikes, lockouts, and other weapons of industrial conflict. However, the manpower shortage in wartime is likely to strengthen the hand of workers in dealing with employers, and this shortage, along with the accompanying need for maximized production, makes any interruptions of production resulting from industrial conflict seem unusually serious and unfortunate. Hence we are likely to depend upon governmental agencies in wartime in an effort to maintain industrial peace and uninterrupted production.

Soon after the entry of the United States into World War II, President Roosevelt called a conference of representatives of labor and industry, at which it was agreed that for the duration of the war there would be no strikes or lockouts, that all disputes between workers and employers would be settled by peaceful means, and that a National War Labor Board should be created for the peaceful settlement of such disputes. The board was set up soon afterward and was charged with adjusting and settling labor disputes that might interrupt work which contributed to the effective prosecution of the war. The board consisted of four public members, four employer members, and four labor members (two representing the American Federation of Labor and two the Congress of Industrial Organizations). Disputes could be turned over to the board by the Department of Labor after its Conciliation Service had failed to settle them, or the board could take jurisdiction over any dispute in a war industry on its own initiative.

In 1943 Congress passed the War Labor Disputes Act, which related to establishments that might be required for the war effort or that might be useful in connection therewith. It authorized government seizure of private plants in which production had been interrupted by strikes, and provided (1) for the continuance of the same working conditions as had prevailed prior to governmental seizure, and (2) for penalties against persons instigating or aiding any strike, slowdown, or interruption in plants taken over by the government. The act also required that thirty days' notice of a

labor dispute in plants covered by it had to be given, and that after the expiration of this waiting period a secret strike ballot should be taken in order to determine whether a majority of the workers favored going on strike. The various measures for maintaining industrial peace in wartime were not completely successful, but there was a considerable reduction in the number and seriousness of work stoppages resulting from industrial conflict. However, we should remember that these measures were only unusual methods for dealing with familiar problems.

Inflation in Wartime. In wartime, prices and the cost of living tend to rise rapidly and spectacularly. The national income, expressed in terms of money paid out to all classes of income receivers, rises to abnormally high levels when a nation produces a huge volume of war goods of all kinds, constructs many new plants for the production of such goods, and continues the production of goods for the civilian population to as great an extent as possible. At the same time, consumers' goods available for purchase are far too limited in quantity to use up all this money income if they continue to sell at prices which were current at the beginning of the war. Unless the money incomes of the citizens are greatly reduced by the payment of taxes and the purchase of government bonds out of current income, price inflation becomes inevitable in the absence of direct governmental control over prices.

The danger of inflation is not peculiar to wartime periods. While extreme inflation often occurs in such periods, other serious and sometimes disastrous inflations have occurred in various countries in peacetime. Moreover, the general causes of inflation are the same in wartime and in peacetime. In either case prices tend to rise when the quantities of money and credit at the disposal of the citizens increase more rapidly than the volume of economic goods for which this purchasing power may be spent. However, the consequences of sharply rising prices in wartime may be considered especially undesirable. For one thing, inflation greatly increases the cost of war in terms of money. World War I cost the United States 31 billion dollars and 13.5 billion dollars of this sum represented inflated prices rather than ships, guns, and other war goods.[2] The rise in prices which occurred from September, 1939, to March, 1942, increased the prospective cost of World War II to the United States by more than the total cost of the preceding war; and it was estimated that, if the price rise were permitted to continue, it would add at least another 50 billion dollars to the cost of the war by the end of 1943.[2]

However, the increase in war costs is only one effect of sharply rising prices. They make it very difficult for businessmen and the government to draw up and carry out plans. If they are to plan and produce effectively,

[2] Office of Price Administration, Consumer Division, Washington, D. C., Releases of April 28 and 29, 1942.

enterprisers must know, within reasonable limits, what their costs and prices will be. Government appropriations and revenue measures are speedily outmoded in periods of sharply rising prices, and war contracts must be revised frequently. Rising prices cause people to buy goods which they do not need at the moment, and to purchase and hoard civilian goods or war materials, either to make a speculative profit or to avoid paying higher prices in the future. The result is interference with orderly and efficient production and a stimulus to further price inflation.

Rising prices impose undeserved sacrifices on persons with fixed or limited incomes, and on members of substandard wage groups whose bargaining power is weak. Rising prices lower the fighting spirit of the people by reducing the purchasing power of their money incomes, and by eating into the value of savings accounts, insurance policies, and bonds. Moreover, there is always the danger that inflation will get wholly out of hand, and eventually reach a point at which money is virtually valueless, the government's credit is destroyed, and production, distribution, and general economic processes are broken down. Finally, an inflationary process, even if it does not reach such disastrous levels, is almost certain to be followed by severe deflation and depression, with millions of persons unemployed, businessmen losing billions on inventories of raw materials and finished goods and failing by the thousands, and farmers suffering bankruptcy and often losing land which they purchased at inflated prices.

Because of such considerations, the government of a nation at war is likely to attempt to control prices directly, whereas in peacetime it may be content to allow prices to find their own level or to exercise only the most general sort of control over the price level by means of monetary and banking policies. In the emergency created by World War II, the United States government, through a succession of agencies culminating in the Office of Price Administration, attempted to control the prices of a number of economic goods. When the United States actually entered the war, Congress soon passed the Emergency Price Control Act in January, 1942. This act continued the powers of price control in the hands of the OPA and its administrator, gave official sanction to the previous activities of the OPA, and made possible the direct governmental control of prices in general.

In April, 1942, the so-called General Maximum Price Regulation was announced. It provided that (1) beginning May 18, 1942, retail prices of commodities, with some exceptions, could not exceed the highest level which each individual seller charged during March, 1942; (2) beginning May 11, 1942, manufacturers' and wholesalers' prices and the prices for wholesale and industrial services could not exceed the highest March levels for each seller; (3) beginning July 1, 1942, no one could legally charge more for services sold at retail in connection with a commodity than he charged during March; (4) effective immediately, all retailers, wholesalers, manu-

facturers, and sellers of services had to preserve existing records of sales made during March, for maximum pricing purposes when the regulation went into effect. The regulation also provided for the immediate licensing of all retailers and wholesalers, effective as of the date when the regulation applied to their particular commodities or services; that is, retailers were directed to regard themselves licensed as of May 18, and wholesalers as of May 11.

The General Maximum Price Regulation, as originally issued, could not be expected to have been completely successful in controlling prices in general, because it did not apply to prices in general. Wages (the prices paid for labor) had been excluded from control, and so also had the prices of basic agricultural commodities, fresh fruits and vegetables, eggs and poultry, butter and other dairy products, fresh fish and other seafood, mutton and lamb, and flour. These excluded commodities and services were important, and their uncontrolled prices continued to rise after the GMPR went into effect. The rises in uncontrolled prices were of significance not only in themselves, but because of the pressure which they brought to bear on the prices of other things. Manufacturers, wholesalers, and retailers found it difficult to keep their maximum prices unchanged in the face of rising costs of labor and agricultural raw materials. It was evident almost from the beginning that price control would have to be extended if it were to operate effectively.

The so-called Anti-Inflation Act of October 2, 1942, was an attempt to deal with some of the pressing problems of price control. In particular, the act laid the foundations for controlling wages and the prices of agricultural products, and delegated broad powers to the president to take action along the lines it indicated. The president at once issued an executive order creating an Office of Economic Stabilization, headed by an Economic Stabilization Director, who was to have important powers in the campaign against inflation. Late in October, 1942, the director issued extensive regulations for the control of wages and salaries. The control of wage payments, and of salaries not in excess of $5,000 per year paid to employees who were represented by a duly recognized or certified labor organization or who were not employed in a bona fide executive, administrative, or professional capacity, was turned over to the National War Labor Board. The control of salaries of more than $5,000 per year, and of other salaries of $30 per week or more paid to individuals employed in executive, administrative, or professional capacities, was given to the Commissioner of Internal Revenue in the Treasury Department.

Under detailed regulations laid down by the National War Labor Board and the Commissioner of Internal Revenue in December, 1942, wages and salaries of $5,000 or less per year were in general frozen at the levels prevailing on September 15, 1942; and salaries higher than $5,000 per year,

or lower salaries paid to individuals employed in executive, administrative, or professional capacities, were frozen at the October 3, 1942, level. Increases in individual wages and salaries of the first type would be permitted only to correct maladjustments and inequalities, to eliminate substandard living conditions, to correct gross inequities, or to aid in the effective prosecution of the war, while salaries of the second type could be increased only to correct maladjustments or inequalities, to aid in the effective prosecution of the war, or in other exceptional cases.

As previously noted, the Anti-Inflation Act also provided for the control of the prices of basic agricultural commodities (such as corn, wheat, tobacco, and cotton). It indicated that maximum prices for such products were to be established at the higher of two levels: (1) the parity prices for the products as determined by the Agricultural Adjustment Administration, and (2) the highest prices received for the commodities between January 1, 1942, and September 15, 1942. On this basis, the president ordered the Economic Stabilization Director, the Secretary of Agriculture, and the Price Administrator to freeze the prices of agricultural products, in so far as it might prove to be practicable, at the levels which prevailed on September 15, 1942. Maximum prices determined in this way exceeded parity for a number of basic agricultural products. Much progress was also made in bringing the prices of other products, such as fresh fruits and vegetables, butter, eggs, poultry, and fresh fish, under control.

Many other changes in the details of price control (as, for example, in methods of computing maximum prices) were made later on, but the direct control of prices in general by the government remained in effect throughout the war period. Beyond saying that this price control operated more effectively after the changes induced by the Anti-Inflation Act were made, it would be beyond the scope of our present study to attempt to evaluate the system of price control in detail or to analyze the many problems which were encountered in its administration. It will suffice for our purposes to note that the whole system of direct price control by the federal government was merely an unusual solution for a familiar problem.

Rationing in Wartime. While consumer spending or purchases as a whole may be quite well maintained in wartime, the great need for commodities necessary to the carrying on of a major war makes it certain that some and perhaps many items of normal civilian consumption will be available in quantities which are much more limited than usual. In such cases we are concerned not only with the prices at which such commodities will sell but with the way in which the limited available stocks are distributed among the consumers. Of course, this problem is peculiar to wartime only in the matter of degree. At all times the available stocks of commodities are too small to permit the complete satisfaction of consumer desires for them, and some way must be found of holding down market

demands for the commodities to the level of the quantities available in the market.

In peacetime we depend upon changes in the prices of the commodities for a solution to this problem. If some commodity is available in unusually limited quantities, a rise in its price will restrict the number of buyers to such an extent that the market demand for and market supply of the commodity will be in equilibrium. Conversely, if the available stock of a commodity is unusually large, a decline in its price will induce an increase in the number of buyers who come forward until the quantity of the good which sellers desire to dispose of is equal to the quantity which buyers are willing to purchase. However, in wartime we may well be unwilling to depend upon this mechanism for the apportionment of consumers' goods among the citizens.

Since many articles are certain to be in unusually short supply in wartime and since relatively large amounts of money and credit are available to consumers as a whole, dependence on the price mechanism for our rationing of scarce commodities would be sure to result in a considerable inflation of prices, with effects which we have noted previously. Moreover, rationing on the basis of price changes distributes scarce commodities on the basis of the size of the money incomes of the individual purchasers and not on the basis of their presumed needs. In wartime, with stocks of certain commodities unusually limited, this might mean that virtually all of some commodities would go to persons with large money incomes, while other individuals would get scarcely any. This might cause considerable hardship and might also affect war production adversely.

Consequently, in wartime we are likely to resort to the rationing of scarce commodities by physical quantities rather than on the basis of price changes. Rationing by physical quantities was applied to a number of commodities in the United States during World War II. Gasoline was rationed by making small allowances to all car operators for ordinary driving and then giving additional quantities to various types of persons for whom extensive transportation by automobile seemed to be in the public interest. In the case of tires, tubes, and new automobiles, no provision at all was made for ordinary civilian uses and the entire available stocks were reserved for essential purposes. Sugar, coffee, and shoes were rationed among individuals by means of ration books and coupons which permitted the purchase of given quantities in given periods of time. Meats, butter, cheese, fats and oils, and canned, frozen, and dried fruits and vegetables were rationed on the basis of other ration coupons and a complicated point system which gave the consumer an element of choice in distributing his purchases among the restricted commodities. However, all this system of rationing by physical quantities represented merely an unusual solution for an essentially familiar problem.

Conclusion. The problems of public finance are also very important in wartime, but there is no need to repeat here any of the material presented in Chapter XXVIII. Our general conclusion is that the economic problems of war are, for the most part, similar in *nature* to those which exist in peacetime, though they may be much greater in *magnitude* and *urgency*. Hence the popular notion that the economic problems of war are something entirely new under the sun is unfounded. The solutions which are provided for economic problems in wartime are usually quite different from those which are relied upon in peacetime and depend to a great extent on direct governmental interference and control. As previously noted, the result is that a capitalistic system is at least temporarily converted into a controlled and planned economy. Actual differences between nominally capitalistic economic systems and those of the socialistic or fascist type are narrowed greatly in time of war.

QUESTIONS AND PROBLEMS

1. "The principle of opportunity costs is as valid in wartime as in peacetime." Explain.
2. "The law of diminishing productivity can hardly be expected to operate under wartime conditions." Show whether you agree.
3. Why are the economic principles of production said to be valid under wartime conditions? Explain.
4. Discuss the status of the principles of saving and capital formation, and those of commercial and investment credit, in wartime.
5. "Some economic principles in the fields of value and distribution are as valid in wartime as in peacetime." Explain.
6. Why are some economic principles, especially in the fields of value and distribution, likely to have diminished practical importance in wartime? Explain.
7. How is the law of comparative advantage affected by wartime conditions? Explain.
8. "The quantity theory of money will operate perfectly in wartime if it is given an opportunity to do so." Do you agree? Explain.
9. How are necessary increases in commodity production accomplished in wartime? Explain.
10. "The conversion and extension of productive facilities are necessary both in wartime and in peacetime, but they are controlled in different ways from the one time to the other." Explain.
11. Why are we likely to resort to direct governmental control over production in wartime? Explain.
12. Explain the scope of manpower or labor problems in wartime. How were these problems handled in the United States during World War II?
13. Why is the problem of inflation especially crucial in wartime? Explain.
14. How were prices controlled by the government in this country during World War II? Explain.

15. "Rationing of finished goods is necessary in both wartime and peacetime." Explain.

16. "The type of rationing used in wartime merely represents an unusual solution for a familiar problem." Do you agree? Explain.

17. "The economic principles and problems of wartime are entirely different from those of peacetime." Show whether you agree.

REFERENCES FOR FUTURE READING

Boulding, K. E., *The Economics of Peace*. New York: Prentice-Hall, Inc., 1945, Chaps. 1–14.

Hardy, C. O., *Wartime Control of Prices*. Washington, D. C.: The Brookings Institution, 1940, Chaps. 1–6.

Harris, S. E. (editor), *Postwar Economic Problems*. New York: McGraw-Hill Book Company, Inc., 1943, Chaps. 1–8, 10–15, 19–23.

Kaplan, A. D. H., *The Liquidation of War Production*. New York: McGraw-Hill Book Company, Inc., 1944, Chaps. 1–5.

Spiegel, H. W., *The Economics of Total War*. New York: D. Appleton-Century Company, 1942, Chaps. 3–10.

Stein, E., and Bachman, J., *War Economics*. New York: Rinehart & Company, Inc., 1942, Chaps. 1–17.

Steiner, G. A., and Associates, *Economic Problems of War*. New York: John Wiley and Sons, Inc., 1942, Chaps. 6–26.

INDEX

Acceptance, discounting of, 467–468
  nature of, 460
Accidents, as workers' risk, 176
Advertising, and human wants, 9, 67
  indivisibilities in, 130
  and monopoly output, 301
  and output under monopolistic competition, 315
  and unproductive activities, 81–82
Agents of production. *See* Productive agents
Aggregate demand, and fiscal policy, 600
  and full-employment policy, 615
  importance of, 598–600
  nature of, 598
Agricultural Adjustment Administration, 149, 557, 620
Agriculture, diminishing productivity in, 108–113
  federal aid to, 553, 557
  and increasing cost, 250
  profits in, 406
  risk in, 175
Alternative costs. *See* Opportunity costs
Anderson, A. G., Mandeville, M. J., and Anderson, J. M., *Industrial Management,* cited, 98
Anti-Inflation Act of 1942, 645–646
Average cost, behavior of, 119, 242
  and competitive price in the intermediate period, 268–270
  and competitive price in the long run, 272–274
  defined, 118–119, 242
  and the enterpriser's return, 253–254
  long-run curves of, 247–248, 302–303
  and marginal cost, 123, 124, 242*n.*, 243*n.*
  and optimum output, 120–121
  and optimum size of firm, 136
  and size of firm, 133–134, 246–248
Average product, curves of, 110–111
  and marginal product, 110–111
  nature of, 108–109
  and the point of diminishing productivity, 110
Average revenue, 228
Average variable cost, behavior of, 118, 240–242
  and competitive output, 125–126, 266–268
  defined, 240
  and marginal cost, 242*n.*
  and monopoly output, 299
Average variable costs, and the point of diminishing productivity, 242

Balance of international payments, favorable and unfavorable, 534–538
  and gold movements, 517–518, 536–537
  maintenance of, 517–519, 521–522, 534–538
Bank checks, as credit instruments, 460–461
  interdistrict clearing of, 473–474
  intradistrict clearing of, 473
  local clearing of, 472–473
  in relation to money, 434
Bank failures, causes of, 499–501
  and deposit insurance, 502–503
  dishonesty and, 499
  and the Federal Reserve System, 498–501
  number of, 499
Bank notes. *See* Federal Reserve bank notes; Federal Reserve notes; National bank notes
Bank reserves. *See* Reserves, bank
Banking, chaps. 24, 25
  central, 474–503
  commercial, 466–474
  investment, 461–466
  See also Commercial banking; Investment banking
Banking Acts of 1933 and 1935, 494–495, 501–503
Banks, Federal Reserve. *See* Federal Reserve Banks
Bargaining power, and distribution theory, 337
  and interest, 372
  and profits, 409–410
  and rent, 349
  and wages, 385
Barter, defined, 194
  difficulties of, 194–195
Basing point systems, 145
Bills of exchange, balancing purchases and sales of, 511–513
  conveniences of using, 509
  defined, 509
  and export financing, 509
  and import financing, 510
  prices of, 513–523
  risks of dealers in, 509–510
  triangular uses of, 511–512

Board of Governors of the Federal Reserve System, *Banking and Monetary Statistics*, cited, 487
Federal Reserve Bulletin, cited, 448, 451, 483, 487, 497, 568
functions of, 485–486
organization of, 485
original powers for credit control, 491–494
present powers for credit control, 494–497
Bonds, characteristics of, 157–158
as claims on wealth, 27–28
as credit instruments, 461
Book accounts, 459–460
Borrowing, public, 579–584
See also Debt, public
Bounties, and international trade, 538
and long-run cost curves, 284
Bullion, gold, 450
silver, 451
Business cycles, and employment, 589
and human values, 589–590
and national income, 588–589
overproduction theory of, 593–594
periods of, 590–593
and price-level changes, 587–588
reality of, 596–598
self-generating theory of, 594–596
See also Economic instability
Business failures in the United States, 586
Business, organization of. *See* Productive unit, organization of
Business risk. *See* Risk
"Buy at Home" theory, criticisms of, 539–540
explained, 539
By-products, and joint products, 286
and optimum size of firm, 129*n.*, 130*n.*

Capital, and consumers' goods, 29–31
control under socialism, 73
controlling total amount of, 63–64, 67
and the corporation, 156–159
defined, 30
distribution among industries, 60, 64
fixed and circulating, 31
free and specialized, 31
and indirect production, 86–87
inequality and formation of, 424–425
and interest, 359–360
as productive agent, 86
and saving, 87–89
scarcity of, 13
supply of, 335
as wealth, 29–31
Capitalization process, 344–345
Certificates, gold, 439, 450–451
silver, 439, 440, 451–452
Chase, Stuart, *Men and Machines*, cited, 102
Checks, bank. *See* Bank checks
Circulating media in the United States, 450–454, 496
Civil Works Administration, 557
Civilian Conservation Corps, 557
Clark, J. M., *The Economics of Overhead Cost*, cited, 132
Class price, and competition, 306
meaning of, 305
and profits, 305–306
Clearing of checks, 472–474
Coinage, described, 437–438
free and gratuitous, 438
as a governmental monopoly, 438
Combinations of productive units, 138–144
complex horizontal, 141–142
nature of, 138–139
number and importance of, 138
simple horizontal, 139–140
success of, 142–144
vertical, 140–141
Commercial banking, chaps. 24, 25
control of credit, 491–498
demand deposits, 468–469
discounting, 466–468
Federal Reserve banks, 484–485
Federal Reserve, Board of Governors, 485–486
Federal Reserve districts, 482
Federal Reserve System, 482–505
functions of, 466–468
interdistrict bank clearings, 473–474
intradistrict bank clearings, 473
loans and discounts, 469
local bank clearings, 472–473
member banks, 482–484
National Banking System, 474–480
problem of elasticity, 470–471, 491–498
problem of safety, 469–470, 497–503
recent changes in, 494–498, 501–503
rediscounting, 486–489
reserves behind deposits, 469–470
time deposits, 468
withdrawal of deposits, 471–472
Commercial credit, and business cycles, 471, 590–593
control under Federal Reserve System, 491–498
and demand deposits, 468–469
and discounting, 466–468
and economic stability, 471
and economic systems, 504–505
elasticity of, 470–471
elasticity under Federal Reserve System, 491–498
expansibility of, 470–471
expansibility under Federal Reserve System, 489–490
expansion in wartime, 496
extent in use, 496
Federal Reserve policy and control of, 493–494
instruments of, 459–461
and investment credit, 458–459
loans and discounts, 469
nature of, 458–459
open market operations, 492–493, 494
rediscount rates, 491–492, 494
reserve requirements, 469–470, 483
Commercial paper, discounting of, 466–468
eligibility of, 487–488
nature of, 486

Commercial paper, rediscounting of, 486–489
Commodity Credit Corporation, 557
Common stocks, characteristics of, 158–159
  as credit instruments, 461
Communism, commercial credit under, 504
  diminishing productivity under, 116
  distribution under, 429–430
  economic institutions of, 54
  indirect production under, 92
  international trade under, 541–542
  investment credit under, 504–505
  large-scale production under, 137–138
  money abandonment under, 454
  risk under, 178–179
  specialization under, 101
  and value theory, 324
Comparative costs, 528–530
Competition, as a capitalistic institution, 45–47
  causes of profits under, 400–407
  and class price, 306
  under communism, 54
  and distribution theory, 336–341
  under fascism, 55–56
  functions of, 46–47
  and interest determination, 371–373
  intermediate-period prices under, 265–270
  intermediate-period supply under, 243–245
  long-run prices under, 272–285
  long-run supply under, 246–254
  as a market condition, 201–202
  and monopolistic competition, 48–49
  and monopoly, 48
  and profits, 400–407
  and rent determination, 348–349
  restrictions on, 47–49
  short-run prices under, 256–265
  short-run supply under, 234–239
  under socialism, 53–54
  summary of price determination under, 284–285
  and theory of international trade, 532–534
  and wage determination, 382–388
Complex horizontal combinations, 141–142
Composite demand, 290
Composite supply, 289–290
Conspicuous consumption, 9–10
  and demand, 216
Constant cost, industries of, 250
  under monopolistic competition, 251, 316–317
  under monopoly, 251, 301
  price determination under, 276–278
  reasons for, 250
Consumers' choices, and control of production, 58–59
  inequality and, 69, 418–419
Consumers' goods, apportionment of, 61–62, 67
  and capital, 29–31
  control under communism, 76
  control under fascism, 76–77
  defined, 30
  rationing under socialism, 74
  scarcity of, 11
Consumers' goods, as wealth, 30–31
Consumers' surplus, 220
Controlled Materials Plan, 640–641
Convertible paper money, 439–440
Cooperation, as a capitalistic institution, 49–50
  under communism, 54
  conscious and unconscious, 49–50
  under socialism, 54
Cooperative associations, 166–169
  consumers' cooperatives, 167
  credit cooperatives, 168
  defined, 166
  importance of, 169
  marketing cooperatives, 168–169
  principles of, 168–169
  producers' cooperatives, 168
Copyrights, as claims on wealth, 27
  and freedom of enterprise, 40
  and monopoly, 204
Corporation, 155–166
  bonds, 157–158
  and capital accumulation, 156–159
  common stock, 158–159
  defined, 155
  dominance of economic activity by, 165–166
  double liability, 160
  importance of, 156–157
  labor relations under, 164–165
  and large-scale production, 155
  limited liability, 159–160
  overcapitalization of, 162–164
  ownership and control of, 161–162
  ownership and management of, 161
  permanence of, 160
  predatory control of, 161–162
  preferred stocks, 158
  and the public, 164–165
  securities of, 157–159
  self-financing of, 602, 607–608
  stock dividends of, 164
  transferability of ownership, 160
Cost, comparative. *See* Comparative costs
Cost of production, average, 118–119, 120–121, 123, 124, 133–134, 242, 247–248, 272–274
  average variable, 118, 125–126, 240–242, 266–268, 299
  differential, 244
  and entrepreneurial income, 253–254
  fixed, 117, 240
  general nature of, 238–239
  industries of constant, 250
  industries of decreasing, 250–251
  industries of increasing, 249–250
  in the intermediate period, 239–243
  and international trade, 532–534
  joint, 285–288
  in the long run, 246–251
  and long-run supply, 251
  marginal, 122–124, 242–244, 248, 265–268, 296–300, 313–315
  opportunity, 238–239
  and prices of finished products, 329

Cost of production, and prices under socialism, 321–323
and profits, 398–399
in the short run, 234–235
and size of firm, 128–134, 246–249
and size of industry, 249–251, 274–276
uniformity of long-run competitive, 251–253
variable, 117–118, 240–242
Credit. *See* Commercial credit; Investment credit
Credit instruments, 459–461
acceptances, 460
book accounts, 459–460
checks, 460–461
drafts, 460
promissory notes, 460
securities, 461
Cyclical fluctuations. *See* Business cycles; Economic instability

Debt, public, 579–584
burden on real income, 581–582
dangers of, 580–584
financial burden of, 582
future burden of, 581–584
growth of, 579–580
inflationary dangers of, 580–581
light-hearted attitude toward, 582
Decreasing cost, false interpretations of, 275–276
industries of, 250–251
under monopolistic competition, 251, 316–317
under monopoly, 251, 301
price determination under, 281–283
reasons for, 250–251
Deductive method in economics, 17–20
Demand, aggregate, 598–600
defined, 212
and differences in desires, 217
and differences in income, 217
diminishing satisfactions, 218*n*.
elastic and inelastic, 222–226
"exceptions" to Law of Demand, 215–217
for foreign exchange, 515–517, 520–521
illustrated, 212–213
illustrations of diminishing utility, 218–219
illustrations of elastic and inelastic, 224–226
increases and decreases in, 213–214
Law of Demand, 215
Law of Diminishing Utility, 218
and marginal revenue, 228–230
and market demand, 220–221
for one grade of labor, 376–379
for one grade of land, 346
for one seller's product, 226–228
popular meanings of, 211
for productive agents, 329–333
real and hypothetical, 215
for savings in consumption, 363–365
for savings in production, 365–366
Demand deposits, elasticity of, 488–489
methods of creating, 468–469
Demand deposits, nature of, 468
withdrawal of, 471–472
Deposit currency. *See* Demand deposits
Deposit insurance, 502–503
Deposits, bank. *See* Demand deposits; Time deposits
Desiredness. *See* Utility
Depression, described, 592
and public expenditures, 557–558
Dewing, A. S., "A Statistical Test of the Success of Consolidations," cited, 142
Dickinson, H. D., *Economics of Socialism*, cited, 50
Diminishing productivity, in agriculture with labor variable, 108–110
in agriculture with land variable, 112–113
assumed conditions of, 114–116
in building, 114
of capital, 365–366
and costs of production, 116–126
and economic systems, 116
illustration of, 108–111
in industry with labor variable, 113
of labor, 376–378
of land, 346
Law of Diminishing Productivity, 114
in merchandising, 113–114
in mining, 114
point of, 110
and practical cultivation, 111–112
and rent, 354–355
in wartime, 631
Diminishing utility, and desiredness, 218–219
Law of Diminishing Utility, 218
and marginal utility, 220
and satisfactions, 218*n*.
Direct production, compared with indirect, 86–87
defined, 86
Discounting, 466–468
Distribution of income, chaps. 17–22
assumptions of theory of, 336–340
capitalization process, 344–345
causes of competitive profits, 400–407
causes of inequality, 425–427
consumptive demand for savings, 363–365
criticisms of inequality, 418–424
defense of inequality, 424–425
demand for one grade of labor, 376–378
demand for one grade of land, 345–346
and demand for productive agents, 329–333
distribution under communism, 429–430
distribution under fascism, 431
distribution under socialism, 428–429
earnings of actual capital, 359–360
as factor pricing, 328
family income distribution, 1929, 415–416
family income distribution, 1935–1936, 416–417
general principles of, chap. 17
into general shares, 327–328
gross and net interest, 360–361
gross and net profits, 398–400
gross and net wages, 388–391
inequality in, 414–428

Distribution of income, interest, chap. 19
interest determination in the long run, 363–373
interest and theoretical assumptions, 371–373
introduction to, chap. 17
long-run supply of labor, 379–381
and marginal product, 330–331
and marginal productivity, 331
and monopolistic competition, 338–340, 349, 372, 387
motives for saving, 367–369
noncapitalistic, 428–431
noncompeting groups of labor, 391–394
personal, chap. 22
personal income distribution, 1944, 417–418
possibility of oversaving, 373–374
productive demand for savings, 365–366
profits and enterpriser efficiency, 401–403
and failures of competition, 409–410
under monopolistic competition, 311–312, 314, 317–319, 407–408
under monopoly, 145–146, 295, 298–301, 303–304, 407
and predatory activities, 408–409
and risks, 403–406
remedies for inequality, 427–428
rent, chap. 18
on all grades of land, 353–354
and assumptions of theory, 348–349
as differential surplus, 352–353
and interest, 359–360
as marginal productivity, 353
for one grade of land, 345–350
and royalties, 344
for several grades of land, 350–354
and unearned income, 355–357
short-run supply of labor, 379
and supply of capital, 334, 365–369
and supply of labor, 334–335, 379–381
and supply of land, 333–334, 345–346
and supply of productive agents, 333–335
supply of savings, 366–369
and time element, 338
time-preference theory, 367–369
uniform net rate of interest, 361–363
and value theory, 329
wages, and theoretical assumptions, 382–388
for several grades of labor, 391–395
for single grade of labor, 376–382
Dividends, on common stocks, 159
as interest, 399
on preferred stock, 158
as profits, 399
stock, 164
Division of Labor. *See* Specialization
Dollar, content of, 433
defined, 433
devaluation of, 447–448
silver, 451–452
Double liability of stockholders, 160
Double-name commercial paper, 467
Drafts, 460
Duopoly, extent of, 205
Duopoly, nature of, 205
price determination under, 306–307

Earnings, reinvested. *See* Reinvested earnings
Economic activity, basis of, 3–14
Economic concepts, chap. 2
capital, 29–30
consumers' goods, 30–31
economic goods, 24
goods, 23
importance of defining, 22
income, 32–36
land, 28–29
problems of defining, 22–23
utility, 23
wealth, 24–28
Economic goods, defined, 24
Economic inequality. *See* Inequality in income distribution
Economic instability, chaps. 29, 30
aggregate demand, 598–600, 615
appraisal of planned economy, 624–627
and business cycles, 590–598
changes in employment, 589
changes in national income, 588–589
changes in price levels, 587–588
compensatory spending, 600
definitions of full employment, 615–617
development of depression, 612–613
effects of booms and depressions, 587–592
full-employment policy, need for, 617–621
and planned economy, 623–624
and private enterprise, 621–623
and human values, 589–590
overproduction theory, 593–594
oversaving and underspending, 611
periods of depression, 592
periods of prosperity, 590–591
periods of recession, 591
periods of recovery, 592–593
question of ultimate causation, 613–615
self-generating theory, 594–596
theory of economic maturity, stated, 600–603
criticized, 603–608
underinvestment or underconsumption, 611–612
underinvestment theory, 599–600, 611–615
Economic institutions, of communism, 54
competition, 45–49
cooperation, 49–50
and distribution theory, 336
economic motivation, 41–43
of fascism, 55–56
freedom of enterprise, 39–41
price system, chap 4
private property, 43–44
of socialism, 50–54
Economic man, 43
Economic maturity, theory of, 600–609
and corporate self-financing, 602, 606–608
criticisms of, 603–608
dangers of, 608–609
and declining population growth, 601, 603–605

Economic maturity, explained, 600–603
and new industries, 601–602, 606
and the vanishing frontier, 601, 605–606
Economic motivation, under communism, 54
described, 41
exceptions to, 42
under fascism, 55
and inequality, 425
restrictions on, 42–43
and social welfare, 42
under socialism, 52–53
Economic optimum, and complex human wants, 70
and economic equilibrium, 67–72
and freedom of consumption choice, 71
and inequality in income distribution, 69
and negative human wants, 71
and unmeasured costs, 69–70
Economic planning, under communism, 75–76
evaluation of, 624–627
under fascism, 76–77
and full-employment policy, 623–624
under socialism, 72–75
Economic self-sufficiency, reasons for, 539, 541
results of, 539–540
Economic Stabilization Director, 645
Economics, bases of, 6–14
deductive method in, 17–20
defined, 6
inductive method in, 16–17
as positive science, 15–16
as science, 14–20
scope of, 14–16
value of, 6
Elasticity of credit, under the Federal Reserve System, 486–498
methods of achieving, 486–490
under the National Banking System, 477–478
nature of, 470–471
original methods of controlling, 491–494
present methods of controlling, 494–495
in wartime, 496–497
Elasticity of demand, definitions, 222–223
illustrations, 223–224
and marginal revenue, 229–230
and types of goods, 224–226
Eligibility of commercial paper, and bank failures, 499–500
described, 487–488
Emergency Housing Program, 557
Endowments of productive agents, and absolute advantage, 528
and comparative advantage, 528–530
and cost of production, 527–528
nature of differences in, 526–527
Enterprisers and economic decisions, 67
and human wants, 10
as labor, 85–86
profits and efficiency of, 401–403
as residual agent of production, 400
risks of, 173–176, 178
wages of, 391–395, 398–399
Equality of opportunity, 41
Equilibrium, economic, and economic optimum, 67–72
and governmental interference, 66
and immobility of productive agents, 64–66
and noncompetitive conditions, 66–67
obstacles to, 64–67
and the price system, 67–68
Exchange, advantages of indirect, 195
difficulties of direct, 194–195
direct and indirect, 194–195
foreign, chap. 26
and specialization, 93
Exchanges, commodity, 182–183
security, 185
Expansibility of credit, nature of, 470–471
under the Federal Reserve System, 489–490
Expansion of the productive unit, chap. 7
combinations, 138–144
and economic systems, 137–138
and indivisibility of productive agents, 128–131
and managerial difficulties, 133
monopolies, 144–149
and optimum size, 133–137
in relation to cost, 128–137
Expenditures, public,
and aid to agriculture, 557
and balanced budget, 580
and changes in general prices, 555
and depression activities, 557–558
and development of highways and schools, 556
and division of labor, 552, 555
evaluation of, 560–563
extravagance in, 555*n*.
federal, 553–554
and governmental functions, 555
growing deficit, 579–580
increases in, 553–554
and increases in population, 554–555
local, 554
and maintenance of capitalism, 562–563, 583–584
and national defense, 553
popular attitude toward, 555*n*.
proper fields of, 559–560
proposed for postwar period, 558–559
for protection, 553
and public debt, 579–580
state, 553–554
test for, 559–560
and unemployment, 557
See also Governmental functions
Exports, correcting excesses of, 517–518, 521–522, 536–537
historical development of, 543–546
invisible, 535–536
under lend-lease, 547–548
of the United States, 1940, 542–543
wartime controls over, 546–548

Failures, bank. *See* Bank failures
Farm Credit Administration, 557
Fascism, commercial credit under, 504
diminishing productivity under, 116

Facism, distribution under, 430
economic institutions of, 55–56
indirect production under, 92
international trade under, 541–542
investment credit under, 504–505
prices under, 76–77
risk under, 178–179
specialization under, 100–101
and value theory, 324
Federal Deposit Insurance Corporation, *Annual Report*, cited, 503
and deposit insurance, 502–503
Federal Power Commission, 558
Federal Reserve Bank Notes, 453–454
Federal Reserve banks, and branches, 482
control of, 484–485
direct loans to member banks, 501
functions of, 485
issuance of credit by, 488–489
issuance of Federal Reserve notes by, 486–488
organization of, 484–485
and rediscounting, 486–489
reserves held by, 487, 489
Federal Reserve Board. *See* Board of Governors of the Federal Reserve System
Federal Reserve notes, amount in circulation, 453
elasticity of, 486–488
expansibility of, 489–490
reserves behind, 452–453
Federal Reserve System, bank failures under, 498–501
Board of Governors of, 485–486
centralization of reserves under, 497
commercial banking and public debt, 496
current banking situation, 496–497
districts of, 482
elasticity of credit under, 488–489
elasticity of currency under, 486–488
eligibility of commercial paper under, 487–488
expansibility of credit under, 490
expansibility of currency under, 489–490
functions of Federal Reserve banks, 485
lack of credit control under, 490–491
new powers for credit control of, 494–495
number of banks in, 483–484
organization of Federal Reserve banks, 484–485
original powers for credit control of, 491–494
powers of Board of Governors of, 485–486
rediscounting in emergencies under, 498–499
requirements for membership in, 482–483
safety for depositors under, 497–503
structure of, 482–486
Fiduciary money, coins as, 437
defined, 437
paper money as, 439–440
representative money as, 439
Financing American Prosperity, cited, 583, 589, 599, 618
Fixed cost, 117, 240, 275
Fixed stock, defined, 233
examples of, 233–234
under monopolistic competition, 312
under monopoly, 295–296
perishable, 262–265
and short-run supply, 234–236
Fixed supply, and competitive price, 262–265
and monopoly price, 293–295
nature of, 262–263
and price under monopolistic competition, 311–312
Florinsky, M. T., *Toward an Understanding of the U.S.S.R.*, cited, 51, 52
Folts, F. E., *Introduction to Industrial Management*, cited, 98
Foreign exchange, chap. 26
conveniences of using, 509
corrective influences under gold standard, 517–519
Corrective influences under paper standard, 521–522
demand for and supply of, 515–517, 520–521
direct uses of, 508–511
and gold movements, 512–513
gold par of, 513–514
and gold shipping points, 514–515
governmental control of, 522–523
premiums and discounts, 516
purchasing power par of, 519–520
rates, defined, 513
under gold standard, 513–519
under paper standard, 519–522
risks of dealers in, 509–510
triangular uses of, 511–513
various rates of, 513
Form utility, creation of, 79–80
Free coinage, 438
Free goods, 24
Free trade, gains from, 531–532
Freedom of enterprise, under communism, 54
defined, 39
and equality of opportunity, 41
under fascism, 55–56
restrictions on, 39–41
and social welfare, 39
under socialism, 51–52
in wartime, 40
Full-bodied money, 437
Full employment, desire for, 615
nature of, 616–618
policy of, 621–624
See also Full-employment policy
Full-employment policy, and aggregate demand, 615
and federal expenditures, 583–584
need for, 617–620
and the planned economy, 623–624
and private enterprise, 621–623
problems of estimation in, 622–623
Future contracts, and hedging, 187–189
and speculation, 183

General Maximum Price Regulation, 644–645
Gold, clause abrogation, 446

Gold, and international obligations, 508
nationalization of, 446
policy of purchasing, 448
price of, 447
reserves of the United States, 448
Gold dollar. *See* Dollar
Gold money, change in bullion content, 447
elimination of payments in, 446
as full-bodied money, 437
profit through devaluation of, 447
withdrawal from circulation, 446
Gold par of exchange, 513–514
Gold Reserve Act of 1934, 446
Gold settlement fund, 473–474
Gold shipping points, and costs of shipping gold, 514
and foreign exchange rates, 514–515
meaning of, 514
Gold standard, abandonment of, 445
advantages of, 443–444
foreign exchange rates under, 513–519
and international stability, 443–444
nature of, 442–443
and present monetary system, 449–450
Gold Standard Act of 1900, 445
Government paper money, 439–441
Governmental enterprises, examples of, 169–170
price principles of, 170
reasons for, 169
Governmental functions, and public expenditures, 555
attitude toward, 555*n*.
efficiency in, 559
expansion of, 555–559
nature of, 552–553
proposed, 558–559
test for, 559–560
Governmental ownership, under capitalism, 44–45, 169–170
competition with private enterprises, 169, 621–622
See also Governmental enterprises
Governmental regulation, of allocation of productive agents, 640–641
of banking, 474–503
and business cycles, 617–621
and competition, 47
of consumption, 646–647
and distribution theory, 337–338
and economic motivation, 42–43
and employment, 641
of foreign exchange, 522–523
and freedom of enterprise, 39–40
and interest, 372
of international trade, 546–548
of labor relations, 642–643
and monopoly, 149
and the price system, 66
of prices, 644–646
and private property, 44
of production, 639–640
and rent, 349
and wages, 385–386, 645–646
in wartime, 639–647
Gratuitous coinage, 438
Greenbacks. *See* United States notes

Hearings on the Revenue Act of 1943, cited, 418
Hedging, betting illustration of, 187
example of, 187–189
by grain-elevator operators, 190–191
limitations of, 190
in manufacturing, 187–189
principle of, 187
and profits, 190
and protection against loss, 187
Holding companies, 138–139

Imperfect competition. *See* Monopolistic competition
Import quotas, 538
Imports, correcting excesses of, 517–518, 521–522, 536–537
historical development of, 543–546
invisible, 535–536
of the United States, 1940, 542–543
in wartime, 548–549
wartime controls over, 546–548
Income, defined, 32
distribution of, chaps. 17–22
and employment, 598–600
money, 34
in prosperity and depression, 588–589
psychic, 34
real and money, 34
of the United States, 35, 589
and wealth, 32–34
Inconvertible paper money, dangers of, 440–441, 580–581
defined, 440
greenbacks as, 452
and inflation, 440–441, 580–581
and reserves, 440
Increasing cost, industries of, 249–250
under monopolistic competition, 316–317
under monopoly, 301
price determination under, 278–281
reasons for, 250
Indirect production, advantages of, 90–91
and capital formation, 87–89
compared with direct, 86–87
defined, 86
and diminishing returns, 91–92
and economic systems, 92
extension of, 91–92
hypothetical development of, 86–87
in modern times, 89–90
and profits, 404
and risk, 173–174
Individual prices, determination of, chaps. 10–16
and profits, 405–406
and speculation, 183
See also Price determination
Indivisibility of productive agents, and cost of production, 128–131
in finance and research, 130–131
general question of, 128–129

Indivisibility of productive agents, machinery and plant, 129
in purchasing and marketing, 130
in the use of labor, 131
Inductive method, in economics, 16–17
Inequality in income distribution, chap. 22
criticisms of, 418–424
defense of, 424–425
and economic insecurity, 421
and economic waste, 420–422
extent of, 1929, 415–416
extent of, 1935–1936, 416–417
extent of, 1944, 417–418
and government, 423–424
and the law, 422–423
and misguidance of production, 418–419
reasons for, 425–427
and total satisfactions, 419–420
Inflation, dangers of, 440–441
effects of, 441, 643–644
and employment creation, 622
and the public debt, 580–581
in wartime, 643–644
wartime control of, 644–646
Inheritance, and inequality, 426
and private property, 43
Instability, economic. *See* Economic instability
Institutions, defined, 39
Institutions, economic. *See* Economic institutions
Insurance, of bank deposits, 502–503
and determinability of loss, 181
and independence of individual risks, 181
premiums and benefits, 180
principles of, 180–181
and reduced losses, 180–181
and reduced uncertainty, 180
and risk, 179–182
social, 182, 421, 558, 583
workman's compensation, 182
Integration, industrial. *See* Vertical combinations
Interest, chap. 19
and capital, 359–360
under communism, 430
consumptive demand for savings, 363–365
under fascism, 430
meaning of, 360
and monopolistic competition, 372–373
motives for saving, 367–369
negative, 370
net and gross, 360–361
oversaving, 373–374, 599–600, 611
productive demand for savings, 365–366
under socialism, 429
supply of savings, 366–369
time-preference analysis, 367–369
uniform net rate of, 361–362
Interlocking directorates, and corporate domination of economic activity, 166
and monopoly, 145
Intermediate period, average cost, 242
cost conditions in, 240–243, 265–266
demand conditions in, 265
fixed cost, 240
Intermediate period, and long-run cost curves, 246–249
marginal cost, 242–243
nature of, 239–240
opportunity cost, 238–239
price determination, competitive, 266–270
under monopoly, 298–301
under monopolistic competition, 313–315
supply of individual competitive seller, 243–244
supply of whole competitive industry, 244–245
variable cost, 240–242
International obligations, and foreign exchange, 508–513
settlement of, chap. 26
and use of gold, 508
and use of money, 508
International trade, chap. 27
absolute advantages in, 528
assumptions of theory of, 526
balance of imports and exports in, 534–537
basis of, 526–530
"Buy at Home" theory, 539–540
comparative advantages in, 528–530
and competitive conditions, 533–534
and economic systems, 541–542
and endowments of productive factors, 526–528
and equilibrium conditions, 532–533
favorable balance of, 534–535
gains from, 531–532
invisible items in, 535–536
and large-scale production, 534
Law of Comparative Costs, 529
and mobility of productive agents, 532–533
and monopolistic competition, 533–534
and monopoly, 533–534
obstacles to, bounties, 538
and "Buy at Home" theory, 539–540
foreign exchange controls, 538–539
import quotas, 538
and infant industries, 540–541
licensing, 538
and preparedness, 541
tariffs, 538
and prices of finished goods, 530
and prices of productive agents, 531
protective tariff, 538
theory of, 526–537
of the United States, 542–549
International trade of the United States, 542–549
balance of payments, 1940, 542
in the future, 549
and international debts, 544–545
invisible items in, 542–543
and lend-lease, 547–548
and the recovery program, 545–546
since 1929, 545–546
and the Smoot-Hawley Tariff, 545
trends in, 543–546
in wartime, 548–549

International trade of the United States, wartime regulation of, 546–549
Investment banking, distribution, 463–464
investigation, 462–463
methods of, 461–464
regulation of, 465
underwriting, 463
Investment credit, and commercial credit, 458–459
and economic systems, 504–505
issuance of, 461–464
nature of, 459
and size of firm, 130–131
sources of, 461–462
volume of, 464–465
in wartime, 633
Invisible imports and exports, nature of, 535–536
significance of, 536
of the United States, 542–543

Joint costs, meaning of, 285–286
price determination under, 287–288
Joint demand, 288–289
Joint supply. *See* Joint costs

"Keep-Money-at-Home" theory, 539–540
Keynes, J. N., *Scope and Method of Political Economy,* cited, 17
Kimball, D. S., *Industrial Economics,* cited, 94
Kimball, D. S., and Kimball D. S., Jr., *Principles of Industrial Organization,* cited, 95, 99

Labor, and business enterprisers, 85–86
business and professional, 392
captains of industry, 392
control under socialism, 73–74
controlling total amount of, 62–63
defined, 84
demand for one grade of, 378–379
diminishing and marginal productivity of, 376–378
distribution among industries, 60
governmental interference and wages of, 385–386, 645–646
immobility of, 64–65
incomes and inequality, 425–426
long-run supply of one grade of, 379–381
noncompeting groups of, 391–396
problems in wartime, 641–643
and population, 379–381
as a productive agent, 84–86
risks of, 176–178
scarcity of, 12–13
semiskilled, 391
short-run supply of one grade of, 379
skilled, 392
standards of living and supply of, 380–381
subsistence and supply of, 379–380
supply of, 334–335, 379–381
and units of labor service, 376–377
unskilled, 391
wages and bargaining power of, 385
and mobility of, 384–385
Labor, wages and bargaining power of, for one grade of, 376–382
and organization of, 382–384
regulation of, 385–386, 645–646
Land, control under socialism, 72
defined, 28–29
and diminishing productivity, 112–113
distribution among industries, 59–60
mobility of, 64
no-rent, 351–352
as part of wealth, 28–29
as productive agent, 28–29
rent on all grades of, 353–354
for one grade of, 349–350
for several grades of, 350–354
scarcity of, 11–12
supply of, 333–334
Large-scale production. *See* Expansion of productive unit; Combinations; Monopolies; Optimum size of firm
Law of Comparative Costs, explained, 529–530
stated, 529
in wartime, 635–636
Law of Demand, 215
Law of Diminishing Productivity, 114
See also Diminishing productivity
Law of Diminishing Utility, 218
Law of Long-Run Competitive Price, 273
Law of One Price, 258
Law of Short-Run Competitive Price, 260
Law of Supply, 236
Legal tender, 438–439
Lend-lease program, 547–548
Leven, M., Moulton, H. G., and Warburton, C., *America's Capacity to Consume,* cited, 35, 416
Liability, of corporation, 159–160
double, 160
limited, 159–160
of partnership, 154–155
or single proprietor, 152
unlimited, 152, 154–155
Licensing of exports and imports, 538
Long run, average cost, 248
changes in productive capacity, 246–249
class price, 305–306
competitive prices under composite demand, 290
under composite supply, 289–290
under constant cost, 276–278
under decreasing cost, 281–283
under increasing cost, 278–281
under joint demand, 288–289
under joint supply, 285–288
cost of production, and competitive price, 272–274
and enterprisers' "profit," 253–254
and size of industry, 249–251, 274–276
and supply, 251
cost curves, 247–249
industries of constant cost, 250
industries of decreasing cost, 250–251
industries of increasing cost, 249–250
interest determination, 353–373

Long run, Law of Long-Run Competitive Price, 273
meaning of, 246
monopoly prices, 301–304
monopoly prices and competitive prices, 304
opportunity costs, 238–239
prices under duopoly, 306–307
prices under monopolistic competition, 316–319
prices under oligopoly, 307–308
rent determination, 345–354
uniformity of competitive costs in, 251–253
wage determination, 376–388
Losses, business. *See* Profits

Management, and expansion of firm, 133
See also Labor
Manufacturing, and creation of form utility, 79–80
diminishing productivity in, 113
hedging in, 187–190
importance of corporations in, 156–157
risk in, 173–175
Marginal cost, and average cost, 123–124, 242*n.*, 243*n.*
and average variable cost, 242*n.*
behavior of, 123–124, 242
defined, 122, 242
in the intermediate period, 242–245
in the long run, 248–249, 302–304
and marginal revenue, 266–270, 298–300, 313–315
Marginal land, 351–352
Marginal product, and average product, 108–110, 330–331
defined, 332
of labor, 377-378
of labor and capital in rent determination, 351
and marginal productivity, 330–331
Marginal productivity, defined, 331
and demand for one grade of labor, 376–379
and demand for one grade of land, 347
and demand for productive agents, 332–333
and demand for savings, 365–366
of labor and inequality, 425–426
and marginal product, 330–331
and monopolistic competition, 338–340
and wages, 331–332, 394
Marginal revenue, and average revenue, 228–229
defined, 228
illustrated, 229
and marginal cost, 266–270, 298–300, 313–315
Marginal utility, 220
Market, nature of, 200–201
Market demand, 220–221
Market sharing, 145
Marketing, and size of firm, 130
Marshall, Alfred, *Principles of Economics*, cited, 274
Maturity, economic. *See* Economic maturity
Member banks, Federal Reserve, and call loans, 502
Member banks, capitalization required of, 483
and direct loans from Federal Reserve banks, 501
failures of, 498–501
and investment affiliates, 501–502
and investment banking, 502
and loans to officers, 502
number of, 483–484
and real estate loans, 484
and rediscounting, 486–489
reserves of, 483
safety for depositors in, 497–503
stock purchase required of, 483
Mergers, 138–139
Mitchell, W. N., *The Organization and Management of Production*, cited, 97
Mobility of productive agents, and distribution of labor, 64–65
and distribution theory, 337
and interest, 362, 372
and rent, 348–349
and theory of international trade, 532–533
and wages, 384–385
Monetary system of the United States, abrogation of gold clause, 446
devaluation of the dollar, 446–448
gold purchasing policy, 449
Gold Reserve Act of 1934, 446
Gold Standard Act of 1900, 445
kinds of money, 450–454
nationalization of gold, 446
nationalization of silver, 449
present monetary system, 449–450
Silver Purchase Act of 1934, 449
silver purchasing policy, 449–450
Money, chap. 23
acceptability of, 434, 436
characteristics of, 436–437
coinage of, 437–438
convertible, 439–440
defined, 434
and economic systems, 454–455
fiduciary, 437
functions of, 435–436
full-bodied, 437
and gold standard, 442–445
as income, 34–35
inconvertible, 440–441
kinds of, United States, 450–454
legal tender, 438–439
paper, 439–441
representative, 439
reserves behind paper, 439–440
value of, 441–442
and wealth, 26–27
Monopolistic competition, and competition, 48
combined with oligopoly, 205, 319
and distribution theory, 338–340
extent of, 208–209
and interest determination, 372
intermediate-period prices under, 312–315
and international trade, 533–534
long-run prices under, 316–318
as market condition, 206–208

Monopolistic competition, nature of, 48, 206–208
and product differentiation, 48, 206–207
and production control, 67
profits under, 316–318, 319, 407–408
and rent determination, 349
as result of patents, copyrights, and trademarks, 204–205
short-run prices under, 310–312
summary of price determination under, 319–320
and wage determination, 387
Monopoly, of an agent of production, 340–341
checks on, 149, 202–203
and competition, 47
defined, 202
and distribution theory, 338–340
and economic progress, 148
evils of, 145–149
extent of, 144, 203–204
and interest determination, 372
intermediate-period prices under, 296–301
and international trade, 533–534
long-run prices under, 301–305
as market condition, 202–204
methods of forming, 144–145
prices and competitive prices, 145–146, 305
and production control, 67
profits of, 146, 298, 303–304, 407
and rent determination, 349
short run prices under, 293–296
and socialism, 323–324
success of, 146
unfair practices of, 146–148, 408–409
and wage determination, 387
Monopsony, 341
Monotony of specialized work, 102–103

National bank notes, advantages of, 475–476
amount in circulation, 453
conditions of issue, 475
inelasticity of, 477
reasons for issuing, 475–476
reserves behind, 453, 475
withdrawal from circulation, 453
National Banking System, bank reserves under, 476
defects of, 476–480
elasticity of credit under, 477–478
elasticity of money under, 477
and national bank notes, 475–477
organization of, 474–475
pyramiding of reserves of, 479
safety for depositors under, 478–480
National Industrial Conference Board, *Economic Almanac,* cited, 553, 554, 576, 580
Enterprise and Social Progress, cited, 29
Mergers in Industry, cited, 142
National Labor Relations Act of 1935, 558–619
National Recovery Administration, 557, 620
National Resources Committee, *Consumer Incomes in the United States,* cited, 416
National Resources Committee, *The Structure of the American Economy,* cited, 166
National War Labor Board, 642
Noncompeting labor groups, 391–395
business and professional workers, 392
captains of industry, 392
criticisms of, 392–393
and labor distribution, 64–65
nature of, 391–392
reasons for, 393–394
semiskilled labor, 391
skilled labor, 392
unskilled labor, 391
wage differences between, 394–395
No-rent land, 351–352
Notes, promissory, defined, 460
discounting of, 466–468

Office of Economic Stabilization, 645–646
Office of Price Administration, 644–645
Oligopoly, crossed with monopolistic competition, 205, 319
nature of, 205
price determination under, 307–308
Oligopsony, 341
Open-market operations, 492–493, 494
Opportunity costs, nature of, 238–239
significance of, 238–239
under socialism, 324
in wartime, 631
Optimum output, 120–121
Optimum size of firm, and cost of production, 128–138
and economic systems, 137–138
and general business conditions, 135
and indivisibility of productive agents, 128–131
and managerial difficulties, 133
practical determination of, 136–137
through time, 135–136
and types of industries, 134–135
Overcapitalization, meaning of, 162
practices resulting in, 162–163
and prices, 163–164
and stock dividends, 164
Overhead costs, 117, 240
Oversaving, 373–374, 599–600, 611

Paper money, advantages of, 439
attributes of, 436–437
convertible, 439–440
functions of reserves behind, 441–442
inconvertible, 440–441
representative, 439
Partnership, advantages of, 153
defined, 153
disadvantages of, 154–155
examples of, 153
significance of, 155
Patent pools, 145
Patents, as claims on wealth, 27
and free enterprise, 40
and monopolistic competition, 204
Philadelphia Evening Bulletin, cited, 419
Place utility, 82

Point of diminishing productivity, 110
Population, growth of, and human wants, 12–13
and labor supply, 379–381
and public expenditures, 554–555
and wage rates, 381
Possession utility, 80–81
Predatory activities, 146–148, 408–409
Preferred stocks, 158
Price, defined, 198
and profit, 404–406
and value, 198–199
Price determination, chaps. 10–16
advertising, and monopoly price, 301
and price under monopolistic competition, 315
assumption of stable price level, 199
average cost, 118–119, 120–121, 123, 124, 133–134, 242, 247–248, 272–274
class price, 305–306
and communism, 324
under competition, in intermediate period, 265–270
in long run, 272–285
in short run, 256–265
summarized, 284–285
competitive market conditions, 201–202
under composite demand, 290
under composite supply, 289–290
constant cost, nature of, 250
price determination under, 276–278
cost of production, adequacy of price equal to, 253–254
average, 118–119, 120–121, 123, 124, 133–134, 242, 247–248, 272–274
constant, 250
decreasing, 250–251
elimination of differential, 251–253
fixed, 117, 240
increasing, 249–250
in the intermediate period, 239–243
joint, 285–288
in the long run, 246–251
long-run changes in, 283–284
marginal, 122–124, 242–244, 248, 265–268, 296–300, 313–315
opportunity, 238–239
and price in the intermediate period, 266–270
and price in the long run, 272–274
and prices under socialism, 321–323
in the short run, 234–235
and size of firm, 128–134, 246–249
and size of industry, 249–251, 274–276
variable, 117–118, 240–242, 266–268, 299
decreasing cost, nature of, 250–251
price determination under, 281–283
demand, actual and hypothetical, 215
composite, 290
elastic and inelastic, 222–226
illustration of, 212–213
increases and decreases in, 213–214
joint, 288–289
Law of Demand, 215
Price determination, demand, Law of Diminishing Utility, 218
and marginal revenue, 228–230
and market demand, 220–221
popular meanings of, 211
for the product of one seller, 226–228
under socialism, 320–321
duopoly, defined, 205
price determination under, 206–207
elasticity of demand, 222–226
extent of competition, 202
extent of monopolistic competition, 208–209
extent of monopoly, 144, 203–204
and fascism, 324
fixed cost, 117, 240
fixed stock, 233–234
fixed supply, 262–263
increasing cost, nature of, 249–250
price determination under, 278–281
interpendent prices, 285–290
intermediate period, nature of, 239–240
under joint costs, 285–288
under joint demand, 288–289
Law of Demand, basis of, 217–220
"exceptions" to, 215–217
stated, 215
Law of Diminishing Utility, illustrations of, 218–219
necessary conditions of, 219–220
stated, 218
Law of Long-Run Competitive Price, explained, 273–276
stated, 273
Law of One Price, 258
Law of Short-Run Competitive Price, illustrated, 260–262
stated, 260
Law of Supply, 236
long run, nature of, 246
marginal cost, 122–124, 242–244, 248, 265–268, 296–300, 313–315
marginal revenue, and elasticity of demand, 229–230
explained, 228–229
illustrated, 229
and marginal cost, 266–270, 298–300, 313–315
market, nature of, 200–201
market demand, 220–221
monopolistic competition, nature of, 48, 206–208
under monopolistic competition, in the intermediate period, 312–315
in the long run, 316–318
in the short run, 310–312
summarized, 319–320
monopolistic market conditions, 202
under monopoly, in the intermediate period, 296–301
in the long run, 301–305
in the short run, 293–296
summarized, 305
monopoly price and competitive price, 145–146, 305

Price determination, oligopoly, defined, 205
price determination under, 307–308, 319
price, defined, 198
product differentiation, 48, 206–207
short run, nature of, 233
under socialism, 320–324
supply, of a competitive industry, intermediate period, 244–245
of a competitive industry, short run, 235–236
of one competitive seller, intermediate period, 243–244
of one competitive seller, short run, 234–235
composite, 289–290
and cost in the intermediate period, 238–245
and cost in the long run, 246–254
and cost in the short run, 233–238
defined, 232
fixed, 262–263
illustration of, 236
increases and decreases in, 236–238
Law of Supply, 236
in the short run, 233–238
value, defined, 198
and price, 198–199
variable cost, 117–118, 240–242, 266–268, 299
in wartime, 633–635
Price leadership, 144
Price-level changes, and business cycles, 587–588
causes of, 441–442
control of, 644–646
and creditors, 587–588
and debtors, 587–588
elimination of, in distribution theory, 336
in value theory, 199
and fixed-income receivers, 587–588
and labor, 587–588
and profits, 404–405, 587
and public expenditures, 555
and Quantity Theory of Money, 442
and salaried workers, 587–588
in wartime, 643–644
Price levels. *See* Price-level changes
Price system, chap. 4
and allocation of productive agents, 59–60, 196
and apportionment of consumers' goods, 61–62, 197
and control of production, 58–59, 196
and economic equilibrium, 67–72
and economic inequality, 69
economic optimum under, 67–68
elimination under communism, 75–76
evaluation of, 67–72
and expression of human wants, 70
functions under capitalism, 58–64
and negative human wants, 71
obstacles to equilibrium under, 64–67
under socialism, 72–75
and total amounts of productive agents, 62–64, 197
Private property, under communism, 54
defined, 43
under fascism, 55
and inequality, 426
restrictions on, 44
and social welfare, 43–44
under socialism, 50–51
Producers' goods. *See* Capital
Product differentiation, 48, 206–207
Production, chaps. 5–9
agents of, 84–86
control, under capitalism, 58–59, 196
under communism, 75–76
under fascism, 76–77
under socialism, 72–74
in wartime, 639–642
defined, 79
and diminishing productivity, 107–116
direct and indirect, 84–92
forms of, 79–81
inequality, and misguidance of, 69, 418–419
and risk, chap. 9
and size of firm, chap. 7
and specialization, 92–104
and types of firms, chap. 8
unproductive activities, 81–84
in wartime, 637–639
Productive agents, alternative uses of, 14
apportionment of, under capitalism, 59–60, 196
under communism, 75–76
under fascism, 76–77
under socialism, 73–74
in wartime, 640–642
controlling total quantities of, 62–64, 197
description of, 84–86
international endowments of, 526–528
international trade, and prices of, 531
proportions of, 107–116
scarcity of, 11–14
Productive factors. *See* Productive agents
Productive units, cooperative associations, 166–169
corporations, 155–166
and economic systems, 170–171
governmental enterprises, 169–170
partnerships, 153–155
single proprietorships, 151–152
size of, chap. 7
types of, chap. 8
Profits, chap. 21
in agriculture, 406
as all residual income, 398–399
average, 410
and business combinations, 142–144
and business cycles, 587
under communism, 430
under competitive conditions, 400–407
defined, 399
and efficiency of enterprisers, 401–403
and failures of competition, 409–410
under fascism, 431
gross and net, 398–400
and hedging, 190

Profits, and individual price changes, 405–407
and industrial expansion, 248–249, 277–278, 279–280, 281–282
and monopolistic competition, 316–318, 319, 407–408
and monopoly, 146, 298, 303–304, 407
as net residual income, 399–400
and predatory activities, 408–409
and price-level changes, 404–405
as reward for risk bearing, 403–404
risks and, 404
and size of firm under monopolistic competition, 319, 407–408
under socialism, 429
and war, 406
Property, private. *See* Private property
Prosperity, 590–591
Protective tariff, 538
Psychic income, defined, 34
and gross wages, 388–390
Public borrowing. *See* Borrowing, public
Public debt. *See* Debt, public
Public expenditures. *See* Expenditures, public
Public finance, chap. 28
Public Utilities Act of 1935, 558
Public Works Administration, 557
Purchasing, and size of firm, 130
Purchasing power par of exchange, and gold par, 520
nature of, 519
reasons for, 519–520
variations around, 520–521

Quantity Theory of Money, explained, 442
in wartime, 636

Rationing, and the price system, 61–62, 197
under socialism, 74
in wartime, 646–647
Recession, 591
Reconstruction Finance Corporation, 557
Recovery, 592–593
Rediscount rate, and credit control, 491–492, 494
Rediscounting, defined, 486
in emergencies, 498
to obtain federal reserve notes, 486–488
to obtain reserves, 488–489
Reinvested earnings, as savings, 367, 462
and stock dividends, 164
Rent, chap. 18
on all grades of land, 353–354
under communism, 429–430
complications in determining, 345
contract and economic, 343
and costs of production, 238–239
as differential surplus, 352–353
and diminishing returns, 354–355
and distribution assumptions, 348–349
under fascism, 430
implicit and explicit, 343
and improvements on land, 343
as marginal productivity, 353
and monopolistic competition, 349
and monopoly, 349
Rent, for one grade of land, 345–350
and purchase price of land, 344–345
reasons for, 345
and risk, 176
and royalties, 344
for several grades of land, 350–354
under socialism, 429
and unearned income, 355–357
and use of land, 343–344
Report of the Secretary of the Treasury, 1928, cited, 553
Representative money, 439
Reserves, bank, centralization of, 497
and credit control, 494–495
of member banks in Federal Reserve System, 483
under National Banking System, 476
problem of, 469–470
pyramiding of, 479
reasons for, 469–470
and rediscounting, 488–489, 498
secondary, 470
Risk, chap. 9
accidents, 176
in agriculture, 175
cyclical unemployment, 177
and economic systems, 178–179
of enterprisers, 173–176, 178
and equilibrium conditions, 404
and hedging, 187–191
and indirect production, 173–174
industrial illness, 176–177
and insurance, 180–182
and interest, 360–361
of landowners and capitalists, 176
and large-scale production, 174–175
modern development of, 173–175
and profits under competition, 403–407
reduction of, 179
and specialization, 175
and speculation, 182–183
technological unemployment, 177
universality of, 173
and wages, 388–389
of workers, 176–178
Robbins, L., *An Essay on the Nature and Significance of Economic Science*, cited, 17
Roundabout production. *See* Indirect production
Royalties, and rents, 344

Safety of bank deposits. *See* Commercial banking; Federal Reserve System
Satisfactions, and economic inequality, 419–420
and the price system, 67–68
Savings, and business cycles, 598–600, 611–612
and capital, under communism, 88–89
under fascism, 88–89
under socialism, 88–89
in wartime, 632
and capital formation, 87–89
excessive, 373–374, 599–600, 611

Savings, inequality and, 424–425
and interest rates, 367–369
motives for, 367
and reinvested earnings, 367, 462
supply of, 366–369
time-preference theory of, 367–369
Scarcity, as characteristic of wealth, 24
of consumers' goods, 11
and economic concepts, 35
of productive agents, 11–13
and rationing of consumers' goods, 61–62
and unemployment, 13–14
Science, defined, 15
economics as, 14–20
positive, 15–16
Securities Act of 1933, 465, 558, 619
Securities and Exchange Commission, 465, 558
Securities Exchange Act of 1944, 465, 558, 619
Selective Service System, 641
Self-generating theory of business cycles, 594–596
Self-interest. *See* Economic motivation
Seligman, E. R. A., *Essays in Taxation,* cited, 563
Service utility, 81
Services, as income, 32–33
and wealth, 25–26
Short run, competitive prices in, 256–265
cost of production in, 234–235
meaning of, 233
monopoly prices in, 293–296
prices under fixed supply, 262–265, 293–295, 311–312
prices under monopolistic competition, 310–312
profits in, 401
rent in, 359–360
return on capital in, 359–360
supply in, 233–238
Silver, as fiduciary money, 437
money in circulation, 451–452
nationalization of, 449
purchasing policy, 448–449
reserves of the United States, 449
Silver Purchase Act of 1934, 448
Simple horizontal combinations, advantages of, 139–140
defined, 139
examples of, 139
Single proprietorship, 151–152
Size of productive unit. *See* Expansion of productive unit; Optimum size of firm
Socialism, apportionment of productive agents under, 73–74
commercial credit under, 504
control of production under, 72–73
criticisms of, 624–627
demand theory under, 320–321
diminishing productivity under, 116
distribution under, 428–429
indirect production under, 92
institutions of, 50–54
interest under, 429
international trade under, 541–542
Socialism, investment credit under, 504–505
large-scale production under, 137–138
money under, 454–455
monopoly price under, 323–324
prices and costs under, 321–322
profits and losses under, 322–323, 429
rationing under, 74
rent under, 429
risk under, 178–179
specialization under, 101
wages under, 428
Soviet Russia, commercial credit in, 504
diminishing productivity in, 116
distribution of income in, 429
economic institutions of, 50–54
indirect production in, 92
international trade of, 541–542
investment credit in, 504–505
large-scale production in, 137–138
money in, 454–455
risk in, 178–179
specialization in, 101
value theory in, 320–324
Specialization, chap. 5
advantages of, 93–94
and business fluctuations, 104
disadvantages of, 101–104
and economic systems, 100–101
by establishments, 97–98
example from automobile production, 95
by geographical areas, 99–100
and indirect production, 93
and large-scale production, 93
and loss of pride in work, 103–104
in management, 97
and manual dexterity, 95
in merchandising, 99
and mobility of labor, 96
and monotony, 102–103
and one-sided development, 104
in the professions, 97
and risk, 175
and saving of time, 96
and simplification of tasks, 96
and standardization, 98
by tasks, 94–95
by trades, 94
types of, 94–99
and unemployment, 101
and use of machinery, 96
Speculation, in commodities, 182–185
evils of, 184–185, 186–187
functions of, 183–184, 185–186
and future contracts, 183
and hedging, 189
nature of, 182–183
and price stabilization, 183–184, 186
and risk, 182–187
in securities, 185–187
Standard of living, and inequality, 420–421
and labor supply, 380–381
Standard money. *See* Full-bodied money
Standardization, 98
Statistical Abstract of the United States, cited, 490, 491, 499, 545, 545, 586

Stigler, G. J., *The Theory of Competitive Price*, cited, 133
Stock dividends, 164
Stock exchanges, regulation of, 465
  speculation on, 185–187
Stocks, as claims on wealth, 27–28
  as credit instruments, 461
  nature of, 158–159
Subsidiary coins, amount in circulation, 452
  as fiduciary money, 437
Subsistence Homestead Projects, 557
Subsistence theory of wages, 379–380
Supply, of capital, 335
  of a competitive industry, in the intermediate period, 244–245
    in the short run, 235–236
  of a competitive seller, in the intermediate period, 243–244
    in the short run, 234–235
  and cost, in the intermediate period, 238–245
    in the long run, 246–254
    in the short run, 233–238
  fixed, 262–263
  of foreign exchange, 515–517, 520–521
  illustration of, 236
  increases and decreases in, 236–238
  of labor, 334–335, 379–381
  of land, 333–334
  Law of Supply, 236
  of one grade of labor, in the long run, 379–381
    in the short run, 379
  of productive agents, 333–336
  of savings, 366–369
  in the short run, 233–238
Survey of Current Business, cited, 35, 135, 136, 170, 548, 549

Tariff Act of 1930, 545–546
Tariffs, as obstacles to trade, 538
  reasons for, 539–541
  results of, 539–540
  and trade of the United States, 545–546
Tax system, certainty, 564
  convenience, 564
  economy, 564
  elasticity, 564
  equity, 564–566
  federal, 567–575
  fiscal adequacy, 564
  local, 576–579
  simplicity, 564
  state, 576–579
  tests of, 564–566
  See also Taxation; Taxes
Taxation, distribution of burden of, 564–566
  incidence of, 566–567
  nature of, 563–564
  opposition to progression in, 566
  principle of ability to pay, 564–565
  principles of, 564
  progressive, 565–566
  proposed changes in, 575
  See also Tax system; Taxes
Taxes, capital stock, 570
  under competition, 573–574
  corporation income, 570–572
  customs duties, 574–575
  direct and indirect, 566
  estate, 572–573
  excess profits, 570–571
  excise, 573–574
  federal, 567–575
  federal estate, 572–573
  federal income, 567–570
  general property, 577–579
  gift, 572
  import duties, 574–575
  inheritance, 576
  land, 578
  license, 576
  local, 576–579
  under monopoly, 574
  motor vehicle, 576
  payroll, 572
  personal income, 567–570
  proposed changes in, 575
  sales, 576–577
  state, 576–579
  state income, 576
  use, 576–577
Technological changes, and diminishing productivity, 115
  and long-run cost tendencies, 283–284
  and risk, 174
  and specialization, 96
  and unemployment, 177
Temporary National Economic Committee, *Competition and Monopoly in American Industry*, cited, 144, 146, 203, 205
  Final Statement of Senator Joseph C. O'Mahoney, cited, 165, 166
  Relative Efficiency of Large, Medium-Sized, and Small Business, cited, 143
  The Structure of Industry, cited, 136, 138, 139, 140, 141
  Taxation, Recovery, and Defense, cited, 157
Terborgh, G., *The Bogey of Economic Maturity*, cited, 603, 605, 606, 607, 608, 612
Thorpe, W. L., *The Integration of Industrial Operations*, cited, 138
Thomas, R. G., *Our Modern Banking and Monetary System*, cited, 449
Time deposits, creation of, 468
  nature of, 468
  withdrawal of, 471–472
Time-preference theory, and consumptive demand for savings, 363–365
  explained, 368
  increasing, 368–369
  and productive demand for savings, 365
  and savings, 367–369
Time utility, 80
Towle, L. W., "Time Deposits and Price Stability," cited, 500
Trade acceptances, 460
Triangular exchange, 511–513

Underconsumption, 599, 611–612
Underinvestment, 599–600, 611, 612
Underwriting, 463
Unearned income, profits as, 400
  and rent, 355–357
Unemployment, and risk, 177
  and scarcity of productive agents, 13–14
  and specialization, 101
  and technological change, 177
United States notes, in circulation, 452
  reserves behind, 452
"Unproductive" activities, alleged types of, 81–83
  criticism of concept of, 81–83
  and the definition of production, 83–84

Value, defined, 198
  and demand and supply forces, 199–200
  and price, 198–199
  and price-level changes, 199
  See also Price determination
Value, theory of, chaps. 10–16
  approaches to, 209
  under communism, 324
  competitive prices, chaps. 13, 14
  under fascism, 324
  introduction to, chap. 10
  nature of demand, chap. 11
  noncompetitive prices, chaps. 15, 16
  supply and cost of production, chap. 12
  under socialism, 320–324
  and the theory of distribution, 329
Variable cost, and average cost, 119–120
  behavior of, 118, 240–242
  defined, 240
  and marginal cost, 242*n*.
  and point of diminishing productivity, 242
  and price in the intermediate period, 266–268, 299
Vertical combinations, advantages of, 139–140
  defined, 139
  examples of, 139

Wages, chap. 20
  and bargaining power, 385
  under communism, 430
  and cost of training, 389
  demand for one grade of labor, 376–379
  and disagreeableness of work, 388–389
  under fascism, 430
  and governmental interference, 385–386
  gross and net, 388–391
  long-run supply of labor, 379–381
  of management, and profits, 398–399
  and marginal productivity, 376–378
  and mobility of labor, 384–385
  and monopolistic competition, 387
  and monopoly, 387
  and noncompeting groups of labor, 391–395
  for one grade of labor, 376–388
  and organization of labor, 382–384
  and psychic income, 389–390
  and risk, 388–389
  for several grades of labor, 391–395
Wages, short-run supply of labor, 379
  under socialism, 428
  and standards of living, 380–381
  and subsistence, 379–380
  and theoretical assumptions, 382–388
  wartime controls over, 645–646
Wantedness. *See* Utility
Wants, advertising and, 10
  biological basis of, 8–9
  changes in, 6–7
  "conspicuous consumption," 9–10
  custom and, 10
  and enterprisers' activities, 10
  expansibility of, 6–7
  as given data, 7–8
  habit and, 10
  imitation and, 9
  specialization and, 10
  as ultimate data, 7–8
War Labor Disputes Act of 1943, 642–643
War Manpower Commission, 641
War Production Board, 640–641
Wealth, capital as, 29–30
  characteristics of, 27
  claims on, 27–28
  consumers' goods as, 30–31
  defined, 24
  and income, 32–33
  individual and social, 27–28
  land as, 28–29
  measurement of, 29
  and money, 26–27
  in relation to services, 25–26
  of the United States, 29
Wootton, B., *Lament for Economics*, cited, 17, 70
Works Projects Administration, 557
World War II, allocation of productive agents in, 640–642
  and basis of economic activity, 630
  control of production in, 639–640
  conversion and extension of facilities in, 639–640
  and diminishing productivity, 631
  exchange controls in, 546–547
  expansion of credit in, 496–497
  industrial conflict in, 642–643
  inflation problems in, 643–646
  international trade during, 548–549
  and the Law of Comparative Cost, 635–636
  and opportunity costs, 631
  price control in, 644–646
  and principles of credit, 632–633
  and principles of price determination, 633–635
  and principles of production, 631–632
  and the Quantity Theory of Money, 636
  rationing in, 646–647
  and restrictions on free enterprise, 40
  and saving and capital formation, 632
  and scarcity of the productive agents, 630
  and the theory of distribution, 635
  and the volume of production, 637–638
  wage control in, 645–646